Disclosure of Mental Health: Philosophical and Psychological Perspectives

Abstract Should people with mental health conditions 'come out proud', disclosing information about their condition(s)? Recent research highlights how disclosing this information can promote empowerment and decrease self-stigma. However, many people with mental health conditions still fear that stigmatization and discrimination will follow if they disclose information about their condition. This paper shows that their fears reflect reality. It shows how recent research on the philosophy and psychology of stereotyping and implicit bias, and how mental states are attributed, suggests that they face a serious dilemma. If they want to avoid being misperceived, for example, as dangerous or incompetent, then both disclosure and non-disclosure bring substantial risks. Disclosure could lead them to be misperceived due to stereotyping. Meanwhile, non-disclosure could lead to their mental states, dispositions, needs and character being misperceived. There are strategies that can be used to reduce the misperception associated with disclosure, i.e. strategies to combat stereotyping. However, whether or not these strategies are implemented is often beyond the control of people with mental health conditions. So, it shall be argued, societal changes are required to ensure that people with mental health conditions can be 'out and proud' without penalization.

Key words: Mental health conditions, stereotyping, implicit bias, mental state ascription, misperception

1. Introduction

People with mental health conditions are often required to address the question of whether they should disclose information about their mental health. Should they inform their employers, colleagues, friends, family, neighbours, etc. that they have a mental health condition? Should they be encouraged by others to do so? There has been a recent move to promote disclosure as a way to increase the empowerment and decrease the self-stigma of people with mental health conditions (see, e.g. Corrigan et al 2010; Corrigan and Rao 2012). For instance, a three week intervention, *Coming Out Proud,* has been devised to inform people about the costs and benefits of disclosure, forms of disclosure, and helpful ways to tell others about one's mental health condition (Rusch et al 2014).

However, many people with mental health conditions continue to fear stigma and penalization if they disclose information about their condition. For example, a recent survey in the UK found that over fifty percent of the mental health service users surveyed feared disclosing their condition because they anticipated stigma and discrimination.[i] In another survey, almost fifty percent of female physicians who were surveyed stated that although they thought that they met the criteria for having a mental illness they had not sought treatment, in many cases due to fear of stigma (Gold et al 2016).

Should people with mental health conditions put these fears aside and respond to the evidence suggesting that disclosure can reduce self-stigma by letting others know about their mental health condition? This paper shows what recent discussions in philosophy can contribute to answering this question. It shows how recent philosophical and psychological work on stereotyping vindicates those who fear stigmatization. I focus in particular on the way that people who have mental health conditions can rightly fear that they will be *misperceived* as a result of others knowing about their mental health condition. In one survey, fifty-one percent of service users said that they wanted the message "We are people – see me, not the illness" to be conveyed because they believed that they were misperceived as a result

of their condition. I show that people with mental health conditions who hold this view are vindicated by recent work in the philosophy of stereotyping: if they disclose information about their mental health condition they face a substantial risk of being stereotyped and consequently misperceived.

It might seem as if the message of this paper is therefore that people should not disclose information about their mental health conditions for fear of stigmatization and being misperceived. However, I show that philosophical and psychological research into how people ascribe mental states to each other suggests that the choice *not* to disclose information about one's mental health condition can also lead to misperception. The main goal of the paper is therefore to outline the contours of a serious dilemma faced by people with mental health conditions: whether or not they disclose information about their mental health, they risk being misperceived.[ii]

2. Disclosure and stereotyping

People with mental health conditions can decide to withhold or disclose information about their conditions. There is legislation to protect them from discrimination if they disclose information. For example, in the United Kingdom the Equalities Act 2010 disability is a protected characteristic and it is stated that "A person (A) discriminates against another (B) if, because of a protected characteristic, A treats B less favourably than A treats or would treat others" (13.1). A mental health condition is considered a disability and discrimination is prohibited. However, legislation cannot be expected to prevent all forms of stereotyping relating to mental health conditions.

Studies have repeatedly shown that in spite of equality legislation, and some reduction to negative attitudes (Crisp 2005; Time to Change 2010), stereotypes remain. A recent UK

Office of National Statistics opinions survey of attitudes towards people with schizophrenia, depression and anxiety uncovered stereotypes associating members of these groups with being a "danger to others", "unpredictable" and "hard to talk to" (Wood et al 2014). These findings replicate earlier studies undertaken by Crisp and colleagues (2000, 2005) and earlier work finding that people with mental health conditions are strongly associated with dangerousness (Link et al 1999; Phelan et al 2000). People with mental health conditions have also been found to be associated with incompetence and character weakness (Corrigan and Watson 2002). There is some variety in ways in which different mental health conditions are stereotyped. Schizophrenia is especially strongly associated with dangerousness and unpredictability (Angermeyer and Matschinger, 2003; Crisp 2000, 2005; Wood et al 2014). Meanwhile, people with anxiety and depression are more likely to be treated as blameworthy for their condition (Wood et al 2014). People with depression are associated with being lazy and not easy to talk to (Thornicroft et al 2007). Further support for the idea that stereotyping frequently occurs comes from the personal experiences of people with mental health conditions. For example, a series of surveys, conducted over the period 2008-2011 in the United Kingdom, found that 87-91% users of mental health services experienced discrimination in four or five areas of their life (Corker et al. 2013).[iii] Discrimination occurs due to stereotyping of individuals being manifest in action.

There is also reason for thinking that many people harbour implicit biases relating to mental illness. Implicit biases are "fast, automatic and difficult to control processes that encode stereotypes and evaluative content, and influence how we think and behave" (Holroyd and Puddifoot forthcoming). Studies have repeatedly found evidence that people with mental illnesses are associated more strongly than others with concepts such as *dangerous, incompetent, threatening,* and *unpredictable* or *helplessness* (Rusch et al 2011; Teachman et al 2006). The finding that people with mental illnesses are associated with negative or

unpleasant concepts has been replicated across various populations: the undergraduate population, the general population, and people with mental illnesses (Teachman et al 2006); medical students with no direct experience of people with mental illness and mental health care professionals (Kopera et al 2015); and people of various age groups including young children and adolescents and older people (O'Driscoll et al. 2012). Various implicit methods have been used to uncover people's implicit biases.[iv] Experimental evidence has been gathered suggesting that implicit biases relating to mental illness manifest in behaviour. The degree of negative implicit bias displayed by mental healthcare professionals has been found to predict their certainty that they would help someone with a mental illness (those with high levels of negative bias said that they were less certain to help) (Brener et al 2013) and the tendency to over-diagnose (those with stronger negative biases were more likely to over-diagnose) (Peris et al 2008).

Recently doubts have been raised about the measures used to identify implicit biases, the stability of implicit biases, the replicability of studies on implicit bias, and about the extent to which implicit biases influence behaviour. Nonetheless, it is worthwhile recognising that implicit biases could operate because if they do so they will exacerbate the problems of stereotyping by leading to unintentional stereotyping by people who are explicitly egalitarian. If people are influenced by implicit bias they may associate people with mental health conditions with characteristics due to their social group membership, thereby stereotyping, without intending to do so. It is especially important to recognise this possibility as research on implicit biases relating to mental illness is well positioned to meet many of the criticisms often levelled at implicit bias research: findings demonstrating implicit biases in this domain have been replicated across different contexts, using various methodologies, and experiments have measured the extent to which implicit bias relating to mental health predicts behaviour.

What is important for current purposes, however, is that there exist stereotypes relating to mental illness, which may operate explicitly or implicitly, and people who have mental health conditions often experience being stereotyped. Legislation aimed at preventing people from discriminating cannot be expected to prevent people from applying stereotypes, e.g. associating people with mental illness with dangerousness. When a person is perceived as belonging to the group of people with a mental health condition, or as having a particular mental health condition, a stereotype associated with the group is likely to be triggered. Consequently, any particular individual with a mental health condition is highly likely to be stereotyped.

3. Stereotyping and misperception

Not only are people with mental health conditions likely to be stereotyped by others who are aware of their condition, they are likely to be *misperceived* as a result of being stereotyped. The philosophical literature on stereotyping has revealed that there are multiple ways that stereotyping can lead to the misperception of individuals to whom a stereotype is applied, whether it occurs implicitly or explicitly (Puddifoot 2017a, b).

When people engage in stereotyping, they often perceive information about individuals in distorted ways that fit the stereotypes that are applied. They notice evidence that is consistent with the stereotype more readily than evidence that is inconsistent with it (Rothbart et al 1979; Cohen 1981; Srull et al 1985; Levinson 2007). The stereotype-consistent evidence is remembered and stereotype-inconsistent evidence is not remembered (ibid). People engaging in stereotyping interpret ambiguous evidence in a way that is consistent with stereotype rather than how it should be interpreted, i.e. as ambiguous (Duncan 1976; Devine 1989). They also explain behaviours in terms of stereotypes where a more accurate explanation would focus on features of the individual case (Sanbonmatsu et al

1994). In addition to this, when stereotypes are activated, and people are viewed through the lens of their social group *and* as members of a minority, it is assumed that they are similar to other members of their minority group (Bartsch and Judd 1993). Differences between members of the same social group are missed. Meanwhile, members of different social groups are assumed to be less similar to each other than they really are (e.g. Tajfel 1981). Similarities between members of different social groups are missed.

With respect to mental health, these phenomena can manifest in the following ways. Evidence about the personal traits or characteristics of someone with a mental health condition that is consistent with them fitting a stereotype—e.g. evidence of their incompetence in some domain—is likely to be noticed and given attention while evidence that challenges the stereotype—e.g. evidence of their competence—is not noticed or attended to. Meanwhile, ambiguous behaviours displayed by people with mental health conditions, such as lateness or inattention—might be viewed as evidence of stereotypical characteristics like having a weak character or being volatile. When an action of this or a similar sort is best explained by the personal traits of the individual, or features of their situation, but not by their mental health condition, the traits are likely to nonetheless be explained in terms of the stereotype: e.g. people with mental illness are like that. Meanwhile, people with mental health conditions, because they are viewed through the lens of their membership of the group, *people with a mental illness*, are likely to be assumed to be similar to other previously encountered people with mental health conditions and dissimilar to others who do not have a mental conditions. Judgements based on these assumptions will not reflect their true similarities and dissimilarities to others.

In addition to this, people with mental health conditions can be denied the opportunity to communicate information about themselves and have uptake of this information from people who engage in stereotyping. Where people have mental health conditions, they can be

treated as if they are unreliable sources of information about their condition when they are actually providing accurate information (Crichton, Carel and Kidd 2016). They can be incorrectly assumed to have mental impairments that mean that they lack self-awareness and cannot be trusted to convey information. This phenomenon is representative of a larger set of phenomena in which people who harbour negative and inaccurate stereotypes can fail to give proper credence to the testimony of those to whom they apply a stereotype (Fricker 2007; Crichton, Carel and Kidd 2016). Where a person believes that their testimony is not likely to receive proper uptake they can choose not to deliver it (Dotson 2011). If this occurs then information about individuals who are stereotyped, information that they might convey through their testimony to others, will not be accessible to those who stereotype them. For example, some information about individual people with mental illness, their personal characteristics and abilities, as well as the symptoms of their conditions, is unlikely to be accessed and utilised when making judgements. If important and relevant information of this sort is not accessed, judgements about people with mental illness are less likely to be accurate or fitting with the evidence that is available.

If a person fears being misperceived due to their mental health condition, their fears sadly reflect reality. As things currently stand, they are likely to be stereotyped due to their mental health condition, and are therefore likely to be misperceived by the person engaging in the stereotyping, who will either fail to access or properly respond to relevant information about them. Disclosure of a mental health condition therefore brings the associated risk of being misperceived, e.g. as dangerous, incompetent, and just like other people with mental health conditions.

4. Non-disclosure and misperception: Information and Understanding

Should people with mental health conditions avoid disclosing information about their condition to avoid the stereotyping and misperception described in sections 2 and 3? The psychological and philosophical literature on stereotyping points in this direction. However, within the philosophy of mind and philosophy of mental disorder, there are other arguments that point in the opposite direction, suggesting that *non*-disclosure of information about a mental health condition can lead to misperception. This is because someone who lacks information about another person's mental health can consequently fail to properly ascribe mental states and dispositions to that person.

For instance, in the philosophical literature on psychiatric delusions, it has been argued that it is possible to understand, predict and explain the behaviour of people with mental health conditions but only if they are viewed as having unusual thoughts. The following example illustrates this point:

> Lizard Man: a 22-year-old Rastafarian man of Jamaican parents was admitted from casualty, having superficially stabbed himself with broken glass. He had become acutely distressed over the past 2–3 days, feeling anxious and depressed and believing that his movements were [being] watched by TV cameras, and signals about him were [being] passed between shopkeepers and that people in shops were talking about him. He was particularly distressed by the scaly appearance of his skin, which he believed was caused by a lizard growing inside his body, the lizard's skin being evident on his arms and legs. He gave the growth of the lizard inside his chest as the reason for stabbing himself. He related this to an incident 10 years before when, in Jamaica, a lizard had run across his face. He believed that the lizard had ''left its mark'' and that a curse then had produced his skin lesions. (Campbell 2009, p. 143, Reimer 2011, p. 662)

Marga Reimer (2011), rightly it seems, claims that it is possible to understand the mental states of the man described in this example. It is even possible to explain why he behaves the way that he does. We could predict related behaviours that he is likely to engage in. However, we are *only* able to do each of these things because we ascribe to him a bizarre thought: that there is a lizard underneath his skin. We need to focus on an aspect of Lizard Man's thought that is so bizarre that the possession of the thought provides a strong indication of mental illness. Without noticing this aspect of his thought, we would not be able to make sense of his actions or the claims that he makes. We would misperceive and misunderstand his behaviour. But on noticing this aspect of his thought it is natural to ascribe him with a mental health condition, i.e. having a psychiatric delusion. It is therefore not possible to avoid misperceiving him and to instead understand what he is thinking without being aware of his mental health condition.

The Lizard Man example is extreme. Many mental health conditions do not include thoughts with such bizarre content. Take, for example, depression, low mood or borderline personality disorder. None of these conditions commonly involves bizarre thoughts. Nonetheless a brief survey of the dominant philosophical accounts of how mental states and dispositions are ascribed shows that the general point about the Lizard Man case applies to these other mental health conditions. In order for people to properly understand certain aspects of what a person is thinking and how they are disposed to think and behave, and in order to accurately predict and explain certain aspects of their behaviour, it will often be necessary to be aware that they have a mental health condition.

Let us begin by considering *simulation theory*. According to this view, when we ascribe mental states to others we project ourselves into their situation (see, e.g. Gordon 1986; Heal 1986; Goldman 1989). We simulate what we would think and feel if we occupied

their situation, making adjustments for any ways that we think that their thoughts are likely to differ from ours. If a person has a mental health condition then they are likely to think and feel in ways that are significantly different to our own. This means that a significant adjustment will need to be made. But the adjustment is only likely to properly reflect the difference between one's own situation and the situation of a person with a mental health condition if one is aware of the condition.

On another account of how mental states and dispositions are ascribed, *theory theory*, mental state ascription involves the application of a tacit theory (e.g Gopnik & Meltzoff 1997; Gopnik et al., 1999). The laws of the theory relate to the situation in which an individual is located, their behaviour as it is found in that situation, and the mental states that underwrite the behaviour. People make inferences from the presence of certain behaviours to the presence of particular mental states by applying the theory. In order for the application of a theory to yield accurate judgements about a person's mental states and dispositions when they have a mental health condition it can be necessary to notice that the tacit theory that successfully predicts many people's mental states and dispositions needs to be adjusted. This is because a person with a mental health condition has aspects of their mental life that differ significantly from most other people's. With respect to those aspects of their mental life, their behaviours will relate differently than other people's to their mental states and situation. For example, the belief that a lizard ran over one's arm and the presence of scaly skin would not usually lead to the behaviour displayed by Lizard Man. A tacit theory that ascribes mental states to people who do not have mental health conditions would be unable to account for Lizard Man's behaviour. In order for Lizard Man's thoughts to be properly understood some adjustment needs to be made to the tacit theory about how mental states, behaviours and situations usually interrelate, to reflect his mental health condition.

Another account of how mental states and dispositions are ascribed is *intentional systems theory*. According to this theory, the ascription of mental states and dispositions is achieved by adopting the *intentional stance*, viewing others as having beliefs, desires, intentions, and so on (Dennett 1971; 1987). The intentional stance involves applying a *principle of charity* to others, assuming that they display rational characteristics; that they believe what they should; have largely true beliefs; and are largely correct in the inferences, expectations and the decisions that they make. People who make these assumptions about people with mental health conditions often draw false conclusions about what they think. In order for accurate judgments to be made about people with mental health conditions it will sometimes be necessary to take their condition into consideration, consequently suspending one's commitment to some or all of the principle of charity, the rationality assumption, and the assumption that they are likely to be displaying true beliefs and correct inferences. Therefore, adjustments will need to be made to the principles underlying most mental state ascription to ensure that the mental states and dispositions of people with mental health conditions are properly understood.

According to a final account of how people ascribe mental states and dispositions to others our capacity for understanding mental states is underwritten by our capacity for creating and applying narratives (Hutto 2008; Gallagher & Hutto 2008). From early childhood, we are exposed to narratives by our caregivers who guide us through the reading of stories such as fairy-tales. In these narratives, characters are introduced within a specific set of circumstances and they act for reasons. Our early exposure to these narratives provides a sense of what is acceptable in certain types of circumstances, what types of activities are interesting and worth noting, what accounts for certain types of action, and what constitutes a good reason for acting. The narratives provide a grasp of cultural norms and likely reasons for action. Then, throughout adulthood, we continue to create and apply narratives to

understand the reasons why individuals act in the ways that they do. On this account of the ascription of mental states and dispositions, it will be necessary to take a person's mental health condition into consideration because the pathological mental states and dispositions of people with mental health conditions will often not fit standard narrative structures. They will violate cultural norms and expectations. To properly understand the mental states and dispositions of those with mental health conditions it will therefore be necessary to factor in how they are likely to deviate from cultural norms and expectations and develop narratives that reflect this.

It is important to note that not all aspects of the mental life of a person with a mental health condition will always be significantly different to the average, deviate from norms of rationality, or violate cultural and social norms and expectations. But there will often be aspects of the mental lives of people with mental health conditions that have these features—the aspects that relate to the mental health conditions. For example, a person with depression might make an unusually negative assessment of their prospects while otherwise having many mental states and dispositions that are like that of the average non-depressed person. It is the assessment of their prospects that means that they have mental states that deviate significantly from the average and it is the same feature of their mental life that means that they are depressed. It is therefore with respect to the aspects of their mental life that are constitutive of their mental health condition, and those closely related, that people risk being misunderstood if they do not disclose information about their mental health.

In sum, then, proper understanding of the mental states and dispositions of people with mental health conditions often requires knowledge about their conditions. The Lizard Man illustrates this point. Each of the dominant accounts of the way that mental states and dispositions are ascribed suggests this to be the case. This means that if a person chooses not to disclose information about their mental health condition, they risk being misperceived by

people who can consequently lack the information required to develop a proper understanding of their mental states and dispositions.

5. Non-disclosure and misperception: What happens in the absence of the information?

Section 3 shows how it can be valuable for a person to disclose information about their mental health: disclosure provides other people with the information required for understanding of their mental states, dispositions and behaviours. This section explores in more depth what happens if they do not disclose the information. What are the specific risks associated with being misperceived due to non-disclosure? This section focuses on two risks, each of which could have a significant impact upon the wellbeing and life outcomes of people with mental health conditions.[v]

First of all, the person who does not disclose information about their mental health condition takes the risk that the needs associated with their mental health condition are not properly understood. Take, for example, a student with anxiety. She dreads public speaking, and therefore misses class when she is aware that she has to orally present her work. If her teacher is not aware of her mental health condition then her absence could be misperceived as indicating that she is lazy, disorganised or lacking commitment to her education. If the teacher misperceives the student in any of these ways then the needs of the student will not be recognised. However, if the teacher has information about the student's mental health condition, special measures could be put in place to support her. For instance, she could be allowed to work as a part of a group, contributing to the research for the presentation without being required to engage in public speaking. There will be many similar cases in which a person has additional needs as a result of their mental health condition. If people are not aware of the condition then their situation will be misperceived and they will not be provided

with the necessary support. And it could have a significant negative impact upon their life outcomes, e.g. leading them to underperform in their education or working life.

Second of all, a person with a mental health condition could be misperceived as having a poor character because of a lack of understanding of their condition. A case that illustrates this point is that of personality disorder. The following is the definition of personality disorder provided by the ICD-10:

> These types of condition comprise deeply ingrained and enduring behaviour patterns, manifesting themselves as inflexible responses to a broad range of personal and social situations. They represent either extreme or significant deviations from the way the average individual in a given culture perceives, thinks, feels, and particularly relates to others. Such behaviour patterns tend to be stable and to encompass multiple domains of behaviour and psychological functioning. They are frequently, but not always, associated with various degrees of subjective distress and problems in social functioning and performance.

If a person displays ingrained and inflexible responses to personal and social situations, which are extreme or significant deviations from the way that the average person behaves in a culture, their behaviour could easily be misperceived. The behaviour can be viewed as indicative of a poor character or 'badness'. Because a person with personality disorder deviates from social norms it could be tempting to describe them as a troublemaker. If, on the other hand, a person discloses that they have a mental health condition, they provide other people with the opportunity to become informed about their condition, and to understand that they are not bad and instead their behaviour is the result of a mental health issue.

Those that do not disclose information about their mental health condition therefore risk being having their needs neglected and being viewed as having a poor character. If they are concerned about being misperceived, then there is good reason for them to be tempted not to disclose information about their condition because doing so reduces the chance of them being misperceived as a result of being stereotyped. However, if they choose not to disclose information about their condition, they risk other forms of misperception. They therefore face a serious dilemma about whether to disclose information about their condition if they want to avoid being misperceived.

6. Objections and replies

It will be worthwhile considering two potential objections that could be raised at this point. Responding to them will provide the opportunity to provide clarification about the nature of the dilemma faced by people with mental health conditions.

6.1. Lack of insight

6.1.1.Objection

The first objection is that people with mental health conditions often lack insight into their condition. Insight is defined as awareness that one has a mental illness, ability to label symptoms of the condition as pathological, and awareness of the effectiveness of and compliance to treatment for one's condition (David 1990). Where people lack insight into their condition, they can lack the ability to disclose information about their condition and would not be likely to face the dilemma presented in this paper.

6.1.2. Reply

It is important to acknowledge that mental health conditions do sometimes involve a lack of insight and that this can provide an impediment to the ability to disclose information about the condition. Where people completely lack awareness that they have a condition it is inappropriate to say that they face a dilemma about whether to disclose that they have the condition. But people with mental health conditions often do have the insight about the nature of their conditions required to disclose information about the condition. Therefore, rather than showing that people with mental health conditions never face the dilemma outlined in sections 1-4, the observation that people sometimes lack insight provides reason to further probe the conditions under which people will face the dilemma.

Recent research suggests that the type of condition that a person has determines the likelihood that they have insight into their condition. For example, a higher number of people with depressive disorder have insight into their condition than people with schizophrenia (e.g. Amador et al 1994; Pini et al 2001; Weiler et al 2000). What this suggests is that whether or not one has the insight required to be able to disclose information about one's condition can depend upon the type of condition it is.

It might be responded at this point that people with delusions, one of the primary cases discussed in section 4, will often fall into the group of people with schizophrenia and the group of people with a mental health condition who lack insight. This could weaken my argument by undermining one of the primary examples used to support it. However, even among the group that has been identified as least commonly having insight—those with schizophrenia—people who are in remission have a good chance of having insight. Take a recent study undertaken by Ramachandran et al. (2016). They measured the insight of patients with schizophrenia and bipolar mood disorder in remission, and found that 60% of patients with schizophrenia were aware of their condition while 100% of patients with bipolar

mood disorder were aware of their condition. Among the group with schizophrenia, 40% lacked insight into their having the condition, but they were in the minority. These results suggest that people with some mental health conditions, such as bipolar disorder, are highly likely to have the insight about their condition required to face the dilemma outlined in this paper when they are in remission. But the results also suggest that many people with conditions often closely associated with a lack of insight, such as schizophrenia, including those with delusions, will also have the required insight when they are in remission. These findings suggest that people frequently face a dilemma about whether to disclose information about their condition not while they symptoms are most acute (see also Jacoby 2016) but in periods when they are less severely ill or in remission. Only then are they likely to have the insight required to disclose information about their condition and therefore the ability to provide information that will allow others to properly interpret their behaviour if they relapse and once again have acute symptoms.

The observation that people with mental health conditions can lack insight does not therefore undermine the argument in sections 1-4. Instead it draws attention to how the likelihood that a person faces the dilemma will depend on factors such as type of mental health condition that they have, and whether they are clinically symptomatic or in remission. The main aim of the paper is not to show that *all* but instead that *many* people with mental health conditions face a dilemma about disclosing information about their condition. Further work on the conditions under which insight occurs has the potential to reveal the conditions under which people with mental health conditions face the dilemma.

It is also useful to note that when a person lacks insight into their mental health condition, others around them can face a dilemma about whether to disclose information about the condition. For example, if one's brother is in acute stage schizophrenia and lacks insight into his condition then one can face a dilemma about whether to disclose this

information, e.g. to his friends. One can fear that he will be misperceived as a result of the stereotyping that might occur if other people being aware of his condition, but if one does not disclose the information, other people are mental states and dispositions. The argument outlined in section 1-4 captures this dilemma that can be faced, for example, by friends and family members, when a person lacks the insight required to disclose information about their condition.

6.2. Un-understandable

6.2.1. Objection

The second objection is that the argument in sections 1-4 is based on the false assumption that disclosure of information about a mental health condition will lead others to understand the mental states of the person with the condition. Using terminology introduced by Karl Jaspers (1913/97), it might be objected that the mental states of people with mental health conditions are *ununderstandable*.

> The most profound distinction in psychic life seems to be between that which is meaningful and *allows empathy* and what in its particular way is ununderstandable, 'mad' in the literal sense, schizophrenic psychic life…" (Jaspers 1913/97, 579)

Jaspers concentrated on people with schizophrenia, but his claim could be applied to other mental health conditions. For example, it might be thought that people with borderline personality disorder are ununderstandable because it is difficult to empathise with them due to the anti-social behaviour that they display. It might be objected that if people with mental health conditions are ununderstandable then disclosure of information about their condition

will not lead to understanding. Therefore, the dilemma outlined in sections 1 to 4 would be illusory. On this view, people with mental health conditions concerned with being properly understood would not gain from disclosing information about their mental health condition because this action would not increase the chance that they would be understood. The choice would then be clear: they should not disclose information because disclosure will lead to stereotyping and misperception.

6.2.2. Reply

There are a number of reasons why this objection does not deal a decisive blow to the claim that people with mental health conditions can face the dilemma outlined in sections 1-4. First, many mental health conditions are easier to understand than schizophrenia and borderline personality disorder. For example, anxiety and depression might be understood by people without the conditions as a result of them extrapolating from their own experiences of being in tense, high pressure situations, or having low mood. Anxiety and depression do not seem to be ununderstandable in the ways that Jaspers identifies. Once again, the central aim of this paper is to show that there is a dilemma faced by *some* and *not all* people with mental health conditions. Therefore, the claim stands as long as there are some people with some mental health conditions whose mental states will only be properly understood if they disclose information about their condition.

However, a number of the most compelling cases of people who have mental health conditions who are required to disclose their condition for others to understand their mental states, e.g. the Lizard man case and borderline personality disorder, are cases in which the conditions are strong candidates for being *un*understandable. Should it be accepted that in these cases disclosure is futile because other people will never be able to understand the person with a mental health condition and their mental states?

There is very good reason to remain more optimistic than this. To see why, consider how it is possible to distinguish between different types of understanding. As Naomi Eilan (2000) argues, on Jaspers' account, understanding involves empathy or what might be called "inside understanding" that involves immersing oneself in another's point of view. Understanding, on this view, involves simulation (ibid.). It is contrast with explanation, which involves knowledge of the causal processes through which mental states emerge. Much of the intuitive force of Jaspers' view derives from the plausibility of the claim that people struggle to immerse themselves in the mental states of those with schizophrenic delusions in the way required to simulate their mental states. However, cases like *Lizard Man* illustrate that it is possible to gain understanding of the mental states of people with psychiatric delusions, by understanding the causal relations between mental states, even if it is not possible to immerse oneself in their perspectives in order to gain inside understanding (ibid.). Even if one cannot imagine what it is like to think that there is a lizard under one's skin one can understand what thoughts and behaviours might arise if one thought this. Cases like this indicate that it is possible to understand the mental states of others, including people with severe mental disorders like schizophrenia, *from the outside* even if they are not understood *from the inside*.[vi]

It will not always be easy to gain this understanding. Merely knowing that someone has, for example, borderline personality disorder will not enable one to understand what they are likely to be thinking, the relationships between their mental states, how they are likely to behave, etc. without some background knowledge about the condition. However, the disclosure of information about one's condition gives other people (family, friends, colleagues, teachers, human resources professionals or other interested parties) the opportunity to find out more about the condition and how it affects thoughts and behaviours.

After research has been undertaken it will be possible for these others to gain understanding from the outside of the mental states and behaviours that are likely to be displayed.

This understanding that can be derived from the outside, without simulating a person's mental states, is valuable and desirable and can prevent a person being misperceived. For instance, if your neighbours are aware that your anti-social behaviour is caused by a delusion that makes you paranoid, they might understand why you behave the way that you do (e.g. your paranoid thoughts lead you to behave suspiciously and act in ways that could be interpreted as rude) and be sympathetic and merciful towards you, even if they cannot imagine what it is like to have the mental health condition that you have. You might prefer that they could fully empathise with you, understanding from the inside what it is like to experience having the condition, but this does not preclude you from valuing the understanding that they can provide.

It remains likely that the mental states of people with mental health conditions sometimes cannot be understood either from the outside or from the inside, or their behaviour predicted or explained. But even in these cases people could benefit from disclosure of the information that they have the condition. If they disclose information other people will be able to make allowances for, excuse, or resist attributions of responsibility for certain behaviours even if they do not understand why the person with the mental health condition thinks and acts precisely as they do.

Let us suppose, as now seems right, that people often lack the ability to fully immerse themselves in the perspective of people with schizophrenia and similar mental health conditions but can nonetheless possess the ability to understand from the outside, or make allowances for, excuse and resist attributions of responsibility. Even those who have mental health conditions with which it is difficult or impossible to empathise, and which would

therefore count as ununderstandable on Jaspers' view, could significantly benefit from disclosure. They therefore face the dilemma outlined in sections 1-4.

7. Solving the Dilemma: Involvement of Others

Once it is acknowledged that many people with mental health conditions face the dilemma outlined in sections 1 to 4, it is natural to ask what can and should be done. A full solution to the dilemma is beyond the scope of the current paper, but this section argues that any adequate solution will require societal change as well as actions by individuals other than those who have mental health conditions. It will then outline some indicative changes. As it will continue to be the case that people with mental health conditions will have aspects of their mental lives that will only be properly understood if their condition is known, the solution to the dilemma will involving ensuring that they can disclose information about their condition without being misperceived due to being stereotyped.

In light of this, it might seem as if people with mental health issues are equipped to respond to the dilemma themselves. They can use the strategy of *selective disclosure*, only disclosing information about their condition to people who they particularly trust (e.g. close family members and friends) or need to tell (e.g. employers), to reduce the chance that people will stereotype them (Bos et al 2009). But there are significant shortcomings of selective disclosure. First, even those people who one particularly trusts may harbour stereotypes or implicit biases relating to mental illnesses. Even close family and friends may be influenced by stereotypes due to their pervasiveness in society, leading them to misperceive the person with the mental health condition. Second, those people who are not told about a person's mental health condition will continue to misperceive them due to their ignorance. Selective disclosure is therefore an imperfect and partial solution to the dilemma.

Alternatives to selective disclosure are found in the form of psychological strategies that can be adopted to change the associations that are made with people with mental health conditions. If a person is aware that they are likely to respond in a biased manner towards members of a particular group, they can adopt long-range strategies to try to change their responses (Holroyd 2012). Long-range strategies are adopted prior to a person being in a situation in which they might be biased. For instance, a person might formulate *implementation intentions* or 'if-then plans', which have been found to be successful in changing the associations that are made with individuals who are encountered (See, e.g. Stewart and Payne 2008). For example, a person might form the plan *if I see a person with a mental health condition then I will think safe*.[vii] Alternatively, one might follow the guidance of advocates of contact theory (Pettigrew and Troop 2005; Corrigan et al 2012; Kolodziej and Johnson 1996; Pinfold et al 2003), who argue that it is possible to change one's stereotyping of stigmatized and marginalized groups through contact with members of the group. One might try to change one's responses towards a certain social group by ensuring that one has regular contact with members of the group, perhaps engaging in a joint task. Other similar strategies are available.

However, it is important to note that the long-range strategies that have just been described are *first personal*; the person whose biases are to be changed must adopt them. This means that, unlike selective disclosure, whether or not they are adopted is beyond the control of the person with the mental health condition(s). For instance, if Ted has a mental health condition, then he might form an implementation intention *if I encounter a person with a mental health condition then I will think safe*. However, Ted's formation of the implementation intention will not prevent other people from stereotyping and consequently misperceiving him. Similarly, Ted cannot decide for other people that they will engage in co-operative activity with people with mental health conditions. Not only do people with mental

health conditions face a dilemma with respect to whether or not to disclose information about their mental health, they are severely limited in their ability to resolve the dilemma. They are highly dependent on others to reduce the extent to which they are misperceived by ensuring that strategies are adopted to combat the stereotyping that leads to misperception.

While this might seem like a pessimistic result, things seem more positive once it is recognised that people with mental health conditions are not wholly dependent upon the willingness of other *individuals* to take action to change the associations that they make. There are changes that can be made to societal structures to reduce the chance of people with mental illnesses being misperceived as a result of stereotyping (see, e.g. Anderson 2010; 2012). It is possible for organisations and institutions to intervene to control how people respond to the information that an individual has a mental health condition, uncoupling mental health conditions from negative associations commonly made with them.

For example, it has been found that although experienced mental health care professionals can display the same amount of negative implicit bias as medical students (Brener et al 2013; Kopera et al 2015), the amount of contact time spent directly with patients in a week predicted the nature of the implicit bias displayed (Dabby et al 2015). This suggests that employers of mental health workers could change people's biases towards those with mental health conditions, decoupling mental health conditions from negative stereotypes, by enacting a policy that requires mental health workers to meet some minimum requirement for the amount of time that they spend each week with service users. In addition to this, in Japan, the old word for "schizophrenia" meaning "split-mind" has been replaced by a new word meaning "integration-disorder". Takahasi et al (2009) found that the old term was strongly associated by participants with "criminal" versus "victim", while the new term was less strongly associated with "criminal". These results suggest that strategies such as the re-naming of mental disorders can be successful at reducing the negative impact of stereotyping.

Meanwhile, anti-stigma campaigners can challenge the associations that are made with mental health conditions. It has been found that emphasising the common group membership (e.g. university attended) between members of different groups (e.g. White vs. Blacks) can reduce stereotyping and increase co-operative activity (for a review see Dovidio and Gaertner 2004). By presenting images of people with mental health conditions that emphasise the features that they share with people without mental health conditions, mental health campaigners can therefore reduce the misperception of people with mental health conditions as dangerous, incompetent, etc. Each of these approaches is piecemeal but governments can enact policies to ensure that initiatives that reduce the negative stereotyping of people with mental illnesses are implemented widely.

There has been a great deal of debate in the recent philosophical literature about whether individualistic strategies that focus on human psychology or broader strategies that focus on social structures are likely to be most effective at reducing inequalities (Banks and Ford 2011; Dixon et al 2012; Haslanger 2015; Huebner 2016, cf. Saul 2013a,b; Madva 2016; Madva 2017). The sensible position seems to be that attempts should be made to change both individuals' psychologies and social structures (see, e.g. Madva 2017). Where there are strategies that can reduce individuals' stereotyping responses there is little reason not to encourage people to adopt the strategies, but wider structural changes can be necessary to ensure that there is wide-ranging equality. For example, the stereotype that people with mental illnesses are incompetent might only be successfully challenged across society if people with mental illnesses have the opportunity to thrive in the workplace, and the contribution that they make is widely recognized. Meanwhile, changing people's psychologies so that they appreciate the need for structural change can be necessary for structural measures to be successfully adopted (Madva 2017). I therefore do not mean to commit to the idea that either individual psychologies or social structures should be the sole

focus of attention for those aiming to reduce the stereotyping of people with mental health conditions. However, changes to social structures, such as institutions, organizations and languages, provide a means of adjusting the responses of many people all at once, without depending on the willingness and motivation of each person. Recognition that these can be changed to reduce the negative impacts of stereotyping of people with mental health conditions therefore provides some reason for optimism.

Conclusion

Should people with mental health conditions 'come out proud', disclosing information about their mental health conditions? Should they be encouraged to do so? This survey of relevant literature from philosophy and psychology highlights a serious dilemma that many people with mental health conditions face with respect to whether or not to disclose the information. If they choose to disclose then they risk being stereotyped, even by those who would explicitly condemn stereotyping. When they are stereotyped they are likely to be misperceived. But if they do not disclose information about their mental health condition then other people are likely to fail to understand their mental states and dispositions, e.g. failing to identify their needs and sometimes judging them to have a poor character because they behave in ways that violate social norms and expectations. People with mental health conditions are not able to tackle this problem on their own. The psychological and philosophical literatures have identified ways to reduce the bias that people face but the person who is likely to be stereotyped cannot enact the methods. Some will only be successful if the person who would otherwise engage in stereotyping adopts them. But changes to the structures of a society, its organizations, institutions and language, can be instrumental in ensuring that stereotype reduction is widespread.

References

Amador XF, Flaum M, Andreasen NC, et al. (1994). Awareness of Illness in Schizophrenia and Schizoaffective and Mood Disorders. *Arch Gen Psychiatry* 51, 10, 826–836.

Anderson, E. (2010). *The Imperative of Integration.* Princeton University Press.

Anderson, E. (2012). Epistemic Justice as a Virtue of Social Institutions. *Social Epistemology*, 26, 2, 163–173.

Banks, R.R. & Ford, R.T (2011). Does Unconscious Bias Matter? *Poverty & Race*, 20, 5, 1-2.

Bartsch R.A. & Judd C.M. (1993). Majority-minority status and perceived ingroup variability revisited. *European Journal of Social Psychology,* 23, 471–83.

Bos, A. E., Kanner, D., Muris, P., Janssen, B., & Mayer, B. (2009). Mental illness stigma and disclosure: Consequences of coming out of the closet. *Issues in Mental Health Nursing*, *30*, 8, 509-513.

Brener, L., Rose, G., von Hippel, C. & Wilson, H. (2013). "Implicit Attitudes, Emotions, and Helping Intentions of Mental Health Workers Toward Their Clients", *The Journal of Nervous and Mental Disease*, 201, 6.

Campbell, J. (2009). What does rationality have to do with psychological causation? Propositional attitudes as mechanisms and as control variables, In M. Broome & L. Bortolotti (eds.) *Psychiatry as cognitive neuroscience: Philosophical perspectives* (pp. 137-149) Oxford: Oxford University Press.

Cohen, C. E. (1981). Personal categories and social perception: Testing some boundaries of the processing effects of prior knowledge. *Journal of Personality and Social Psychology,* 40, 441-452.

Corrigan, P.W., Morris, S., Larson, J., Rafacz, J., Wassel, A., Michaels, P., Wilkniss, S., Batia, K. & Rüsch, N. (2010). Self-stigma and coming out about one's mental illness. *Journal of community psychology* 38, 3, 259-275.

Corrigan, P. W., Morris, S. B., Michaels, P. J., Rafacz, J. D., & Rüsch, N. (2012). Challenging the public stigma of mental illness: a meta-analysis of outcome studies. *Psychiatric services*, *63*, 10, 963-973.

Corrigan, P. W., & Rao, D. (2012). On the self-stigma of mental illness: Stages, disclosure, and strategies for change. *The Canadian Journal of Psychiatry*, *57*, 8, 464-469.

Corrigan, P. W., & Watson, A. C. (2002). Understanding the impact of stigma on people with mental illness. *World psychiatry*, 1, 1, 16-20.

Crichton, P., Carel, H. and Kidd, I.J. (2016). Epistemic Injustice in Psychiatry, *BJPysch Bulletin*, 41, 2, 65-70.

Dabby, L., Tranulis, C., & Kirmayer, L. J. (2015). Explicit and Implicit Attitudes of Canadian Psychiatrists Toward People With Mental Illness. *The Canadian Journal of Psychiatry*, *60*, 10, 451-459.

David, A.S. (1990). Insight and psychosis. *The British Journal of Psychiatry*, 156, 6, 798-808.

Dennett, D. (1971). Intentional Systems. *Journal of Philosophy*, 68: 87–106.

Dennett, D. (1987). True Believers: The Intentional Strategy and Why it Works. In D. Dennett, *The Intentional Stance* (14-35) Cambridge MA: MIT Press.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, 56, 5–18.

Dixon, J., Levine, M., Reicher, S. & Durrheim, K. (2012). Beyond Prejudice: Are Negative Evaluations the Problem and Is Getting Us to Like One Another More the Solution? *Behavioral and Brain Sciences*, 35(6), 411–425.

Dotson, K. (2011) 'Tracking epistemic violence, tracking patterns of silencing,' *Hypatia*, 26, 2, pp. 236–257.

Duncan, B. L. (1976). Differential social perception and attribution of intergroup violence: Testing the lower limits of stereotyping of blacks. *Journal of Personality and Social Psychology*, 34, 590-598.

Eilan, N. (2000). On understanding schizophrenia. In D. Zahavi (ed.) *Exploring the self*: *Philosophical and psychopathological perspectives on self-experience* (John Benjamins Company), 97-113.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing.* Oxford: Oxford University Press.

Gallagher S. & Hutto, D.D. (2008). Understanding Others Through Primary Interaction and Narrative Practice. In J. Zlatev, T. Racine, C. Sinha & E. Itkonen (eds.), *The Shared Mind: Perspectives on Intersubjectivity* (John Benjamins Publishing Company).

Gendler, T. S. (2011). On the Epistemic Costs of Implicit Bias. *Philosophical Studies*, 156, 1, 33-63.

Gold, K. J., Andrew, L. B., Goldman, E. B., & Schwenk, T. L. (2016). "I would never want to have a mental health diagnosis on my record": A survey of female physicians on mental health diagnosis, treatment, and reporting. *General hospital psychiatry*, *43*, 51-57.

Goldman, A. (1989). Interpretation Psychologized. *Mind and Language*, 4: 161–85.

Gopnik, A. & Meltzoff, A.N. (1997). *Words, Thoughts and Theories* (Cambridge, MA: MIT Press).

Gopnik, A., Meltzoff, A.N. & Kuhl, P. (1999) *The Scientist in the Crib* (New York: Harper Collins).

Gordon, R., (1986). Folk Psychology as Simulation. *Mind and Language*, 1, 158–71.

Haslanger, S. (2015). Social Structure, Narrative, and Explanation. *Canadian Journal of Philosophy*, 45, 1, 1–15.

Heal, J. (1986). Replication and Functionalism. In J Butterfield (ed.) *Language, Mind & Logic*, (135-150) Cambridge: Cambridge University Press.

Holroyd, J. (2012), Responsibility for Implicit Bias. *Journal of Social Philosophy,* 43: 274–306.

Holroyd, J. (2015). Implicit bias, awareness and imperfect cognitions. *Consciousness and cognition*, *33*, 511-523.

Holroyd, J. & Puddifoot, K. (forthcoming). Implicit Bias and Prejudice. In M. Fricker, P.J. Graham, D. Henderson, N. Pedersen, and J. Wyatt (ed.) *Routledge Handbook of Social Epistemology*

Huebner, B. (2016). Implicit Bias, Reinforcement Learning, and Scaffolded Moral Cognition. In M. Brownstein and J. Saul (eds.), *Implicit Bias and Philosophy: Metaphysics and Epistemology* (Vol. 1, 47–79). Oxford University Press.

Hutto, D., (2008). *Folk Psychological Narratives: The Sociocultural Basis of Understanding Reasons* (Cambridge MA: MIT Press).

Kolodziej, M. E., & Johnson, B. T. (1996). Interpersonal contact and acceptance of persons with psychiatric disorders: a research synthesis. *Journal of consulting and clinical psychology*, *64,* 6, 1387

Kopera, M., Suszek, H., Bonar, E., Myszka, M., Gmaj, B., Ilgen, M., Wojnar, M. (2015). Evaluating Explicit and Implicit Stigma of Mental Illness in Mental Health Professionals and Medical Students. *Community Mental Health Journal*, 628-634.

Lai, C. K., Marini, M., Lehr, S. A., Cerruti, C., Shin, J. E. L., Joy-Gaba, J. A., ... & Frazier, R. S. (2014). Reducing implicit racial preferences: I. A comparative investigation of 17 interventions. *Journal of Experimental Psychology: General*, *143*, 4, 1765.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., ... & Simon, S. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*,8, 1001.

Lassiter, C., & Ballantyne, N. (2017). Implicit racial bias and epistemic pessimism. *Philosophical Psychology*, *30*, 1-2, 79-101.

Levinson, J. D. (2007). Forgotten racial equality: Implicit bias, decision-making, and misremembering. *Duke Law Journal*, 345-424.

Link, B. G., Phelan, J. C., Bresnahan, M., Stueve, A., & Pescosolido, B. A. (1999). Public conceptions of mental illness: labels, causes, dangerousness, and social distance. *American journal of public health*, *89*, 9, 1328-1333.

Madva, A. (2016). A plea for Anti-Anti-Individualism: how oversimple psychology misleads social policy. *Ergo, an Open Access Journal of Philosophy*, *3*.

Madva, A. (2016). Virtue, Social Knowledge and Implicit Bias. In M. Brownstein & J. Saul (eds.) *Implicit Bias and Philosophy, Volume 1: Metaphysics and Epistemology*. Oxford: Oxford University Press.

O'Driscoll, C. O., Heary, C., Hennessy, E., & McKeague, L. (2012). Explicit and implicit stigma towards peers with mental health problems in childhood and adolescence. *Journal of Child Psychology and Psychiatry*, 53, 1054–1062.

Peris, T. S., Teachman, B. A., & Nosek, B. A. (2008). Implicit and explicit stigma of mental illness: Links to clinical care. *Journal of Nervous and Mental Disease*, 196, 752–760.

Pettigrew, T. F., & Tropp, L. R. (2005). Allport's intergroup contact hypothesis: Its history and influence. In J. F. Dovidio, P. Glick, & L. A. Rudman (Eds.), *On the nature of prejudice. Fifty years after Allport*. Oxford: Blackwell.

Phelan, J., Link, B., Stueve, A., & Pescosolido, B. (2000). Public conceptions of mental illness in 1950 and 1996: What is mental illness and is it to be feared?. *Journal of Health and Social Behavior, 41,2*, 188-207.

Pinfold, V., Toulmin, H., Thornicroft, G., Huxley, P., Farmer, P., & Graham, T. (2003). Reducing psychiatric stigma and discrimination: evaluation of educational interventions in UK secondary schools. *The British Journal of Psychiatry*, *182*, 4, 342-346.

Pini, S., Cassano, G. B., Dell'Osso, L., & Amador, X. F. (2001). Insight into illness in schizophrenia, schizoaffective disorder, and mood disorders with psychotic features. *American journal of psychiatry*, *158*(1), 122-125.

Puddifoot, K. (2017a). Stereotyping: The Multifactorial View. *Philosophical Topics*, *45*, 1, 137-156.

Puddifoot, K. (2017b). Dissolving the epistemic/ethical dilemma over implicit bias. *Philosophical Explorations*, *20*(sup1), 73-93.

Reimer, M. (2011). A Davidsonian Perspective on Psychiatric Delusions. *Philosophical Psychology,* 24, 5, 659-677.

Rothbart, M., Evans, M. & Fulero, S. (1979). Recall for Confirming Events: Memory Processes and the Maintenance of Social Stereotypes. *Journal of Experimental Social Psychology*, 15, 343-55

Rüsch N, Corrigan PW, Todd A, Bodenhausen GV (2011). Automatic stereotyping against people with schizophrenia, schizoaffective and affective disorders. *Psychiatry Res.* 18, 34-39.

Rüsch, N., Abbruzzese, E., Hagedorn, E., Hartenhauer, D., Kaufmann, I., Curschellas, J., Ventling, S. et al. (2014). Efficacy of Coming Out Proud to reduce stigma's impact among people with mental illness: pilot randomised controlled trial. *The British journal of psychiatry* 204, 5, 391-397.

Sanbonmatsu D.M., Akimoto S.A. & Gibson B.D. (1994). Stereotype-based blocking in social explanation. *Personality and Social Psychology Bulletin,* 20, 71–81.

Saul, J. (2013a). Scepticism and Implicit Bias. *Disputatio*, 5, 37, 243-263.

Saul, J. (2013b) .Implicit bias, stereotype threat, and women in philosophy. In F. Hutchison & F Jenkins (ed). *Women in philosophy: What needs to change* (39-60) Oxford: Oxford University Press.

Srull, T. D., Lichtenstein, M., & Rothbart, M. (1985). Associative storage and retrieval processes in person memory. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11, 316-345.

Stewart, B. D., & Payne, B. K. (2008). Bringing automatic stereotyping under control: Implementation intentions as efficient means of thought control. *Personality and Social Psychology Bulletin*, *34*, 10, 1332-1345.

Tajfel, H. (1981). *Human Groups and Social Categories: Studies in Social Psychology* (Cambridge: Cambridge University Press).

Takahashi, H., Ideno, T., Okubo, S., Matsui, H., Takemura, K., Matsuura, M., ... & Okubo, Y. (2009). Impact of changing the Japanese term for "schizophrenia" for reasons of stereotypical beliefs of schizophrenia in Japanese youth. *Schizophrenia research*, *112*(1), 149-152.

Teachman, B. A., Wilson, J. G., & Komarovskaya, I. (2006). Implicit and explicit stigma of mental illness in diagnosed and healthy samples. Journal of Social and Clinical Psychology, 25, 75–95

Weiler, M. A., Fleisher, M. H., & McArthur-Campbell, D. (2000). Insight and symptom change in schizophrenia and other disorders. *Schizophrenia Research*, *45*(1-2), 29-36.

[i] https://www.time-to-change.org.uk/sites/default/files/Stigma%20Shout.pdf

[ii] In the existing literature on the philosophy of implicit bias there has been a debate about whether people face a dilemma with respect to their own implicit biases relating to race and crime in the United States. According to Gendler (2011), if people are influenced by their biases then they can make various errors, but if they are not

influenced by their biases then they will commit base-rate, failing to acknowledge the statistical reality that Black people engage in criminal activity at a higher rate than White people (cf. Madva 2016; Lassiter and Bannatyne 2017; Puddifoot 2017a, b). The dilemma discussed in the current paper is distinct from Gendler's dilemma. I discuss a dilemma that people face with regards to information that they can disclose that can lead others to be biased in their perception of them. The dilemma I identify relates to the biases of other people whereas Gendler's dilemma relates to how one could respond to one's own biases.

[iii] There was some improvement over the period covered by the study (number of people who had experienced discrimination was 91% in 2008, 87% in 2009 and 2010, and 88% in 2011), during which the Time for Change campaign, which aimed to tackle prejudice and discrimination towards people with mental health issues, was in operation. However, the improvement was not statistically significant, nor did it meet the target of reduction of experiences of discrimination set for the campaign. (Corker et al. 2013)

[iv] The Implicit Association Task (IAT) (Teachman et al. 2006); a lexical priming task (Rusch et al. 2011), and the GNAT task (Kopera et al. 2015).

[v] This section aims to provide an illustrative rather than a complete list of the misperceptions that can occur.

[vi] It is consistent with a number of the views outlined in section 4 that understanding *from the outside* can occur. The application of, and adjustment to, tacit theories, principles of charity and narratives to understand the mental states of people with mental health conditions can occur even while one struggles to experience the mental states *from the inside*. Therefore, on many of the dominant accounts of mental state attribution, mental states of people with schizophrenic delusions could be *un*understandable *from the inside*, but could nonetheless be understandable, through the application of a theory, principles of charity or narratives.

On simulation theory, things are more complicated because the simulation of mental states involves empathy. If it is not possible to empathise with people with mental health conditions like schizophrenia then understanding will be impossible. One might conclude that it is not possible to understand the mental states of people with schizophrenia. However, evidence that we can understand the mental states of people like Lizard Man counts against this conclusion, suggesting that either Jaspers is incorrect or simulation theory cannot account for the full breadth of understanding that people can have of the mental states of others.

[vii] Two recent rigorous studies have called into question the effectiveness of long term strategies that have been proposed to tackle implicit stereotyping (Lai et al 2014; Lai et al 2018). In the first study is was found that many were ineffective in the short term (implementation intentions were effective), and where the strategies were effective the efffectiveness was short-lived. These studies provide additional reason for thinking that societal changes will be needed to tackle stereotyping, at least the implicit type.