# Mental disorders and the notion of harm

**Biturajac, Mia**

UNIVERSITY OF RIJEKA
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
DEPARTMENT OF PHILOSOPHY

Mia Biturajac

# MENTAL DISORDERS AND THE NOTION OF HARM

DOCTORAL THESIS

Rijeka, 2022.

UNIVERSITY OF RIJEKA
FACULTY OF HUMANITIES AND SOCIAL SCIENCES
DEPARTMENT OF PHILOSOPHY

Mia Biturajac

# MENTAL DISORDERS AND THE NOTION OF HARM

DOCTORAL THESIS

Advisor: Luca Malatesti, Full Professor
Co-advisor: Marko Jurjako, Associate Professor

Rijeka, 2022.

SVEUČILIŠTE U RIJECI
FILOZOFSKI FAKULTET
ODSJEK ZA FILOZOFIJU

Mia Biturajac

# MENTALNI POREMEĆAJI I POJAM ŠTETE

DOKTORSKI RAD

Mentor: prof. dr. sc. Luca Malatesti
Komentor: izv. prof. dr. sc. Marko Jurjako

Rijeka, 2022.

Mentor rada: prof. dr. sc. Luca Malatesti
Komentor: izv. dr. sc. Marko Jurjako

Doktorski rad obranjen je dana _____ u/na _____, pred povjerenstvom u sastavu:

1. _____
2. _____
3. _____
4. _____

**ACKNOWLEDGEMENTS**

I am grateful to my mentors Dr. Luca Malatesti and Dr. Marko Jurjako for their guidance and valuable advice. They have consistently facilitated my growth and learning through challenges and by upholding my work to the highest standards.

I thank the Croatian Science Foundation (HRZZ) for funding my doctoral research through the *Young researchers' career development project - Training of new doctoral students (Grant: DOK-2018-09-5165)* and *Responding to antisocial personalities in a democratic society (Grant: HRZZ-IP-2018-01-3518)* project.

I am grateful to friends and colleagues from the philosophy department in Rijeka for showing me kindness, patience, guidance and support in my philosophical as well as personal explorations - including but not limited to Dr. Boran Berčić, Dr. Iris Vidmar Jovanović and Dr. David Grčki.

Special thanks to Dr. Ana Gavran Miloš who opened my eyes to the wonders of ancient philosophy and who appeased my insatiable philosophical curiosity of that area. Thanks also to Dr. Majda Trobok. From a person who hated math and could not understand it, she turned me into a person that very much likes math and logic!

I thank Dr.² Ivan Cerovac for not only making me a better philosopher but for making me a better, kinder and a more thoughtful person. His unwavering friendship and support has gotten me through the worst of times and it has brought about the best of times. Here's to our virtue friendship!

I would like to thank my family – parents Mira and Vlado, brother Sven, and my grandparents. This would not be possible without their encouragement and support.

I have also been blessed with amazing friends who have helped me throughout this whole process. You know who you are.

And then, there is Tanja, who has always been there for me.

**SUMMARY**

In this thesis I investigate the notion of harm in psychiatry. The question is whether a person is harmed, and in what way, by a particular condition that is of psychiatric relevance. In other words, I explore whether harm should be one of the criteria to focus on in psychiatric practice - in the process of psychiatric assessment, diagnosis, and treatment of patients. If so, how should we think about, define and conceptualize harm to fit the needs of this context.

What harm in psychiatry entails becomes a bit clearer once we consider how it entered into the psychiatric context. The notion of harm entered psychiatry in the 1970s during the process of depathologisation of homosexuality – the removal of homosexuality from the list of mental disorders. The argument was that mental disorders are conditions that harm a person that has them. Seeing as homosexuals are not harmed in virtue of their condition, they are not suffering from a mental disorder. In years following the APA's decision to depathologize homosexuality, harm would be included in authoritative diagnostic manuals and would be featured in the discussion on mental disorders, which has been a central discussion in philosophy of psychiatry for the last forty years.

In this work I argue that harm is an important criterion in psychiatry, rightfully deserving of our attention. Harm is relevant both in theory and practice as one of the criteria that seriously considers the patient's point of view as we have seen in the case of depathologizing homosexuality.

Even though the traditional debate on mental disorders presents the conceptual and methodological landscape of our analysis, in recent years the debate has been strongly objected to. The lion's share of the criticism has been directed towards traditional philosophical methods like conceptual analysis. Further objections are directed at the methodology of the discussion on mental disorders,

questioning the nature of the concept, the intuitions and goals of the discussion and whether we are investigating one concept or many.

I take these objections to be justified and side with metatheoretical innovations in the discussion on mental disorders that have come about in recent years (Bortolotti, 2020; Ereshefsky, 2009; Murphy, 2006; Schwartz, 2007a). To place harm in the context of these methodological novelties in philosophy of psychiatry I take Dominic Murphy's (2006) taxonomy that builds on the traditional positions in the discussion, objectivism and constructivism, by adding the variants – conservativist and revisionist. My contribution to the discussion is an addition to Murphy's taxonomy, a position which I call "compassionatism". It is the view that concepts that capture the first-person perspective are and should remain crucial in psychiatry. This includes concepts such as harm, well-being, welfare, suffering and so on. I believe that in highlighting compassionist concepts and making them salient we are bringing attention to a key element of psychiatric practice – the patient's perspective. I focus on harm as one of these compassionatist concepts. I take the revisionist route of compassionatism in arguing that our analysis of harm should be constructive and normative, rather than a descriptive project. We should decide on what we need and how we should think about harm rather than describe how harm is used in various psychiatric contexts. Our compassionatist concepts, specifically harm, should be constructed according to two principles. First, the concept should as fully and as faithfully capture the patient's perspective. Second, it should do so while adhering to the precepts of the psychiatric practice.

To come to the concept of harm I use an explicationist methodology inspired by Rudolf Carnap (1962). The method consists in taking an imprecise concept (explicandum) and turning it into a precise concept (explicatum) which would fit particular needs and interests. To begin I use the idea of harm in bodily medicine which, I argue, consists in the inability to achieve and maintain homeostasis. I try to

transpose this view to psychiatry, being also wary of the differences between them and challenges that such a move would pose. I propose what I consider to be a close approximate of this view in psychiatry. I argue that harm in psychiatry consists in inadequate resources in dealing with the problems of living. I unpack each of the elements of the view – 'inadequate', 'resources', 'dealing with' and 'problems of living' showing what they mean and how they work together.

**Keywords:** harm, mental disorder, conceptual analysis, philosophical explication, medicine, philosophy, psychiatry, resources, problems of living

**SAŽETAK**

U ovom radu se bavim pojmom štete u psihijatriji. Pitanje kojim se bavim je šteti li, i na koji način, osobi određeno stanje koje je od psihijatrijskog interesa. Drugim riječima, istražujem bi li pojam štete trebao biti jedan od kriterija na koji bismo trebali obratiti pažnju unutar psihijatrijske prakse – tijekom procesa psihijatrijskog vrednovanja, dijagnoze i liječenja pacijenata. Ako pretpostavimo da jest, kako bismo trebali promišljati, definirati i konceptualizirati pojam štete da odgovara potrebama psihijatrijskog konteksta.

Značenje pojma štete u psihijatriji je jasnije ako razmotrimo okolnosti u kojima je pojam štete uveden u psihijatriju. Pojam štete je uveden u psihijatriju 1970tih tijekom procesa depatologizacije homoseksualnosti – uklanjanja homoseksualnosti s liste mentalnih poremećaja. Argumentom korištenim u tu svrhu se tvrdilo da su mentalni poremećaji stanja koja štete osobama koje boluju od njih. Budući da homoseksualnost sama po sebi ne šteti osobama koje su homoseksualne orijentacije, takvim osobe ne pate od mentalnog poremećaja. U godinama koje su slijedile nakon odluke Američkog psihijatrijskog društva o depatologizaciji homoseksualnosti, pojam štete ulazi u autoritativne dijagnostičke priručnike i proteže se kroz raspravu o mentalnim poremećajima koja je temeljna rasprava u području filozofije psihijatrije posljednjih četrdeset godina.

U ovom radu tvrdim da je pojam štete bitan kriterij u psihijatriji koji s pravom zaslužuje našu pažnju. Pojam štete je bitan i u teorijskoj i u praktičnoj psihijatriji kao jedan od kriteriji koji ozbiljno uzimaju u obzir perspektivu pacijenta, kao što smo to vidjeli na primjeru depatologizacije homoseksualnosti.

Iako je tradicionalna rasprava o mentalnim poremećajima konceptualni i metodološki temelj naše analize, posljednjih godina dobiva mnoge kritike. Najviše kritika je upućeno tradicionalnim filozofskim metodama poput konceptualne analize. Daljnji

prigovori su usmjereni metodologiji rasprave o mentalnim poremećajima koji propituju prirodu pojma, položaj intuicija u raspravi i ciljeve rasprave, te postavljaju pitanje istražuje li se jedan pojam ili više njih.

Smatram da su ti prigovori opravdani i priklanjam se metateorijskim inovacijama u raspravi o mentalnim poremećajima koje su iznesene posljednjih godina. Kako bih smjestila pojam štete u kontekst tih metodoloških novina u filozofiji psihijatrije oslanjam se na taksonomiju pozicija Dominica Murphyja. Njegova taksonomija je izgrađena na pozicijama tradicionalne rasprave – objektivizmu i konsktruktivizmu kojima Murphy pridružuje dvije varijante – konzervativizam i revizionizam. Moj doprinos raspravi je dodatak Murphyjevoj taksonomiji, poziciju koju nazivam 'suosjećajizam'. To je gledište prema kojemu pojmovi koji utjelovljuju perspektivu pacijenta jesu i trebali bi biti važni u psihijatriji. To uključuje pojmove poput štete, dobrobiti, blagostanja, patnje i druge. Stava sam da ističući pojmove suosjećajizma i naglašavajući ih pridajemo pažnju ključnom elementu psihijatrijske prakse – perspektivi pacijenta. U svojoj analizi se usredotočujem na pojam štete koji je jedan od tih pojmova. Zagovaram revizionistički suosjećajizam čime tvrdim da naša analiza štete treba biti normativna, a ne deskriptivna, što znači da bismo trebali moći odlučiti što trebamo od pojma štete te kako bismo trebali promišljati o njemu radije nego da opisujemo kako se pojam koristi u raznim psihijatrijskim kontekstima. Naši 'suosjećajistički' pojmovi, posebice šteta, bi trebali biti konstruirani prema dva principa. Prvo, pojam bi trebao što potpunije i vjernije zahvatiti perspektivu pojedinca. Drugo, to bi trebao činiti poštivajući pravila i prilike psihijatrijske prakse.

Do pojma štete dolazim koristeći metodu eksplikacije inspiriranu Rudolfom Carnapom. Metoda se sastoji u pretvaranju nepreciznog pojma (explicanduma) u precizniji (explicatum) koji bi odgovarao pretpostavljenim potrebama i interesima. Za početak koristim generalnu ideju štete u medicini koja se sastoji u nemogućnosti postizanja i održavanja homeostaze. Nastojim to gledište štete prenijeti u područje

psihijatrije uzimajući u obzir razlike i izazove koje takav potez podrazumijeva. Predlažem definiciju koju smatram bliskom navedenoj definiciji u medicini. Definiram pojam štete kao neadekvatne resurse pri nošenju s životnim problemima. Svaki element te definicije pobliže objašnjavam te pokazujem kako svi elementi funkcioniraju zajedno.

**Ključne riječi:** šteta, mentalni poremećaji, konceptualna analiza, filozofska eksplikacija, medicina, psihijatrija, filozofija, resursi, problemi življenja

**CONTENTS**

# Introduction

In this thesis I investigate the notion of harm in psychiatry. The question is whether a person is harmed, and in what way, by a particular condition that is of psychiatric relevance. In other words, I explore whether harm should be one of the criteria to focus on in psychiatric practice - in the process of psychiatric assessment, diagnosis, and treatment of patients. If so, how should we think about, define and conceptualize harm to fit the needs of this context.

What harm in psychiatry entails becomes a bit clearer once we consider how it entered into the psychiatric context. The notion of harm entered psychiatry in the 1970s during the process of depathologization of homosexuality – the removal of homosexuality from the list of mental disorders. The argument was that mental disorders are conditions that harm a person that has them. Seeing as homosexuals are not harmed in virtue of their condition, they are not suffering from a mental disorder. In years following the APA's decision to depathologize homosexuality, harm would be included in authoritative diagnostic manuals and would be featured in the discussion on mental disorders, which has been a central discussion in philosophy of psychiatry for the last forty years.

In this work I argue that harm is an important criterion in psychiatry, rightfully deserving of our attention. Harm is relevant both in theory and practice as one of the criteria that seriously considers the patient's point of view as we have seen in the case of depathologizing homosexuality.

Even though the traditional debate on mental disorders presents the conceptual and methodological landscape of our analysis, in recent years the debate has been strongly objected to. The lion's share of the criticism has been directed towards traditional philosophical methods like conceptual analysis. Further objections are directed at the methodology of the discussion on mental disorders, questioning the nature of the concept, the intuitions and goals of the discussion and whether we are investigating one concept or many.

I take these objections to be justified and side with metatheoretical innovations in the discussion on mental disorders that have come about in recent years (Bortolotti, 2020; Ereshefsky, 2009; Murphy, 2006; Schwartz, 2007a). To place harm in the context of these methodological novelties in philosophy of psychiatry I take Dominic Murphy's (2006) taxonomy that builds on the traditional positions in the discussion, objectivism and constructivism, by adding the variants – conservativist and revisionist. My contribution to the discussion is an addition to Murphy's taxonomy, a position which I call "compassionatism". It is the view that concepts that capture the first-person perspective are and should remain crucial in psychiatry. This includes concepts such as harm, well-being, welfare, suffering and so on. I focus on harm as one of these compassionatist concepts. I take the revisionist route of compassionatism in arguing that our analysis of harm should be constructive and normative, rather than a descriptive project. We should decide on what we need and how we should think about harm rather than describe how harm is used in various psychiatric contexts. Our compassionatist concepts, specifically harm, should be constructed according to two principles. First, the concept should as fully and as faithfully capture the patient's perspective. Second, it should do so while adhering to the precepts of the psychiatric practice.

To come to the concept of harm I use an explicationist methodology inspired by Rudolf Carnap (1962). The method consists in taking an imprecise concept

(explicandum) and turning it into a precise concept (explicatum) which would fit particular needs and interests. To begin I use the idea of harm in bodily medicine which, I argue, consists in the inability to achieve, and maintain homeostasis. I try to transpose this view to psychiatry, being also wary of the differences between them and challenges that such a move would pose. I propose what I consider to be a close approximate of this view in psychiatry. I argue that harm in psychiatry consists in inadequate resources in dealing with the problems of living. I unpack each of the elements of the view – 'inadequate', 'resources', 'dealing with' and 'problems of living' showing what they mean and how they work together.

The first chapter of this thesis provides an overview of harm in psychiatry. The first part of the chapter tells the story of introduction of harm into psychiatry. It chronicles the case of depathologization of homosexuality through which I show how depathologization of homosexuality was the result of three important elements of psychiatry coming together – the patient perspective, scientific studies in psychiatry and psychiatric/clinical expertise. The second part of the chapter outlines much of the instances of harm in psychiatry found in the literature. I first discuss harm in diagnostic manuals – the Diagnostic and Statistical Manual (DSM) and the International Classification of Diseases (ICD) which are esteemed and widely adopted manuals for psychiatric assessment. Then, the chapter turns to harm in philosophy of psychiatry, in the debate of mental disorder where I discuss and evaluate various accounts of harm in the discussion.

The second chapter deals with the mental disorder debate which is central to philosophy of psychiatry as well as to the notion of harm. It is necessary to investigate harm on the backdrop of the mental disorder debate since it is connected to it terminologically, methodologically, and contextually. First, I go through the role and uses of the notion of mental disorder in various contexts to outline its purpose in psychiatry. I lay out the debate on mental disorders in philosophy of psychiatry with

their positions of naturalism, normativism and hybrid accounts. I point to strengths and weaknesses of each of these positions. I put focus on how harm relates to each of these positions.

In the third chapter I outline the metaphilosophical ideas that have gained traction in recent and contemporary philosophy, especially the philosophy of psychiatry. The recent climate criticizes the traditional debate on mental disorder and the philosophical method of conceptual analyses. We see an abundance of novel approaches that try to remedy some of the issues that are facing the traditional debate. The criticism is directed at conceptual analysis, the nature of the concepts involved, the problem of intuitions in the debate on mental disorders as well as to the number of concepts that we may be dealing with in the traditional discussion. I conclude the chapter in a climate that is reformative with regards to the traditional precepts of the debate on mental disorders.

In the fourth chapter I deal with eliminativism with regards to the notion of mental disorders. As I have pointed out in the third chapter, many methodological and theoretical novelties arose as a response to the purported stalemate of the traditional debate and the numerous problems that can be ascribed to it. One of the responses to such problems is to eliminate the notion of mental disorder from psychiatry which is the idea behind eliminativism. However, I do not argue for nor against eliminativism. Instead, I argue that even if we adopt eliminativism, harm still has its rightful place in psychiatric practice. I end the chapter in reconsideration of eliminating the mental disorder as it is a somewhat controversial position which might cause more problems than it would solve. Granted, I do remain agnostic with respect to eliminating the notion of mental disorder.

The fifth chapter is dedicated to outlining the meta-debate on mental disorder. As the third and fourth chapter have shown that there are problems with the traditional debate on mental disorders as well as with its methodology, the fifth

chapter introduces and lays out new and the (methodologically) improved debate on mental disorders that is heavily packed with metaphilosophical elements. In this I follow Dominic Murphy's taxonomy of the positions in the meta-debate – objectivism and constructivism, each of which can be either conservative or revisionist. Instead of arguing for either of the positions, I offer an addition to the Murphy's taxonomy in the form of compassionatism. Compassionatism consists in bringing to the forefront the concepts that aspire to take into account the patient's perspective – harm, welfare, well-being, suffering and so on. I argue that these concepts are and should remain relevant in psychiatric practice. I advocate for revisionist compassionatism and outline the principles that we should adhere to in devising our compassionist concepts.

In the sixth chapter I set the groundwork for building an account of harm in psychiatry. I propose the methodology of such a project that relies on Rudolf Carnap's method of philosophical explication. As a starting point I use Peter Strawson's idea of conceptual analysis on which I build the process of philosophical explication. Strawson proposes that in our conceptual analysis we should strive to make our implicit understanding of the concepts explicit. Thus, as a starting point I make explicit the idea of harm in bodily medicine. I argue that in bodily medicine harm consists in the inability of an organism to achieve and maintain homeostasis. My goal is to try to translate this view to the domain of psychiatry. However, there are several problems in doing so since there are crucial dissimilarities between bodily medicine and psychiatry which I outline in the last section of the sixth chapter. All of these serve as a starting point in constructing the account of harm in psychiatry.

I dedicate the seventh chapter to devising an account of harm. I propose that harm consists in inadequate resources in dealing with the problems of living. I outline each of these elements and show how they connect to various psychological studies and other concepts that broaden and deepen our understanding of harm in

psychiatry. I dedicate the last section to argue that harm constructed this way is compatible with an idea of well-being proposed by Michael Bishop that defines well-being as positive causal networks. His ideas are evocative of the literature on psychological resources which I have used in constructing the account of harm.

# 1. Harm in Psychiatry

## 1.1 Introduction

On the 15th of December 1973, the American Psychiatric Association declared homosexuality would no longer be considered a mental disorder.

Up to that point homosexuality was thought of as a sin, harshly condemned by the Judeo-Christian social climate. It was a crime prohibited and punishable by law. And it was a mental disorder, one with numerous other comorbidities observed in the clinical setting. All of that was about to change in 1950s and the decades following the first robust scientific investigations on homosexuality. At that time, we also witness the uprise of the gay rights movement and their claim for recognition. Finally, we see the facilitation of the dialogue between the gay community and professional psychiatry. As a result of tenacious activist work and the sympathetic ear of the American Psychiatric Association, homosexuality ceases to be a mental disorder.

In the throes of depathologizing homosexuality we see the emergence of harm as a criterion in defining and diagnosing mental disorder. The question that started to assume importance was whether a particular condition harms the person that has it. This was an unprecedented move in psychiatry as it opened the door for evaluating conditions from the first-person perspective and taking seriously into account the patient's point of view. On the basis of such move, harm as a criterion gained prominence and homosexuality was depathologized and excluded from the diagnostic manuals. It was argued that for a condition to be a disorder it had to harm the person

that has it. Since homosexuality is not intrinsically harmful – many can live happy and healthy lives despite their condition, homosexuality is not a mental disorder.

This thesis is about the notion of harm in psychiatry. The aim of this thesis is to outline and analyze the different ideas and accounts of harm in psychiatry as well as the methodology with which harm is approached. Since harm is rarely a stand-alone criterion in psychiatry it is instructive to analyse harm inside the context it is usually found – the diagnostic manuals (DSM, ICD) and in the debate on mental disorder in philosophy of psychiatry.

Since it was introduced and gained importance in the light of depathologizing homosexuality I take this as a case study of harm in psychiatry. I extract an important lesson from the study in depathologizing homosexuality. On the account of it I propose that there are three pillars of good psychiatric practice – scientific research, patient's perspective and psychiatric/clinical expertise, as these were the elements that were all made to work together in order to come to depathologizing homosexuality. I dedicate this chapter to exploring the case of depathologizing homosexuality and the important lessons that we can extract from that monumental phenomenon.

### ***1.1.1*** *Scientific Research on Homosexuality*

The importance of scientific research in depathologizing homosexuality should not be understated, although at that time it often was. Evelyn Hooker and her research especially should be highlighted as she might have been the first to tackle the issues of homosexuality with such dedication, scientific fortitude and the depth of immersion she was willing to forego in understanding the phenomenon. Inspired by her former student Sam Fromm who urged her to do studies on himself and his homosexual friends, she set out on a journey of uncovering and understanding homosexuality that lasted for several decades between 1950s and 1970s (Herek & Garnets, 2015). Her work was especially laudable as it was done in the time of the

McCarthy era in the U.S when homosexuals and communists, as well as their 'accessories' were endangered and targets of persecution, victims of a "destructive witch hunt" (Hooker, 1993, p. 450). Nonetheless, Evelyn Hooker was unyielding in her exploration as she conducted countless interviews with homosexuals at her own estate. It was an effort to provide a cloak of secrecy and confidentiality to her interlocutors, being conscious of the 'triple stigma' of homosexuality, it being "a sin, a crime, and a disease"(Hooker, 1993, p. 451).

Her research however did not stop at the individual interviews in her home-setting. She immersed herself into the gay community exploring and observing the multitude of its aspects as she recounts:

> Even then, had I chosen just to remain in my study and let the gay men come to me, perhaps the stress would have been less. I could not settle for less. Instead, I accepted invitations to gay parties, gay organizations, gay after-hours clubs, and gay bars. I was convinced that, because of the secrecy imposed on gay men whose occupations and very lives were at risk if their identity became known, it was essential to know and understand everything I could about the gay social milieu that they created. (Hooker, 1993, p. 451)

Her scientific contribution to the issue of homosexuality starts with a 'Preliminary Analysis of Group Behavior of Homosexuals' (1956) where she lays out some remarks and observations she has made in her explorations of homosexuals described above. She starts off by asserting that the prevailing theories of homosexuality fail to consider that homosexuals are starting to view themselves as a minority group, "sharing many of the problems of other minority groups, having to

struggle for their “rights” against the prejudices of a dominant heterosexual majority” (Hooker, 1956, p. 217).

Drawing on Allport’s influential work *The Nature of Prejudice* (1954) she uses the conceptual apparatus of minorities and groups to analyse some of the characteristics of homosexuals. She proposes that some of the characteristics of homosexuals might not be inherent but could be “a) responses of minority or out-group members to cultural pressures of the dominant majority” or “( b ) responses to the in-group pressures of the homosexual groups once the individual accepts membership in them.” (Hooker, 1956, p. 224). They point to the idea that some of their behaviour, even if presumed to be unusual or pathological might not be intrinsic to homosexuality. For example, as Hooker postulates “obsessive concern” with their own sexuality and “withdrawal and passivity” might serve as a kind of “ego defense” against the unfavourable, oftentimes cruel social circumstances they reside in (Hooker, 1956, p. 219). Furthermore, some of the behaviour might also be attributed to their belonging to the gay community with its own rules of conduct, ways and peculiarities (Hooker, 1956). Thus, she poses an important question that directs her further inquiry:

> Are they [passivity and dependence] inherent in the personality structure of homosexuals, or are they in part the result of attempting to cope with a hostile world in a manner which may, in the mind of the victim, produce the least damage to himself? (Hooker, 1956, p. 219)(brackets added)

This served as a prima facie reason for re-examining the pathological status of homosexuality. Hooker would offer answers to the aforementioned question in her most influential piece ‘The Adjustment of the Male Overt Homosexual’ (Hooker, 1957). In it she matched 30 male homosexuals without other pathological conditions,

to 30 male heterosexuals in age, education and IQ. The materials used to study them "consisted of a battery of projective techniques, attitude scales, and intensive life history interviews." (Hooker, 1957, p. 20), which also included Rorschach, TAT and MAPS tests (Hooker, 1957). To the psychologists of this day and age the use of the Rorschach[1] test might seem problematic because of its highly contestable validity, however, Rorschach was used by many clinicians of the time in assessing the personality structure. It was also the test used for diagnosing homosexuality. Besides wanting to obtain 'unbiased judgment' of the personality structure of the individuals, she also wanted to verify "the accuracy which expert clinicians who are Rorschach workers can differentiate homosexual from heterosexual records" (Hooker, 1957, p. 21). The results were that neither of the judges did better than chance in determining whether an individual was homosexual or not, with other tests – MAPS and TAT offering much of the similar results (Hooker, 1957, p. 21). The results pointed out that homosexual males were quite indistinguishable from heterosexual males on various psychological measures. Hooker (1957, p. 29) acknowledges in her paper that:

> what is difficult to accept (for most clinicians) is that some homosexuals may be very ordinary individuals, indistinguishable, except in sexual pattern, from ordinary individuals who are heterosexual. Or-and I do not know whether this would be more or less difficult to accept-that some may be quite superior individuals, not only devoid of pathology (unless one insists that homosexuality itself is a sign of pathology) but also functioning at a superior level.

[1] Rorschach inkblot test is a "projective method of psychological testing in which a person is asked to describe what he or she sees in 10 inkblots, of which some are black and grey and others have patches of color" (taken from https://www.britannica.com/science/Rorschach-Test June 30th 2022)

She concludes her work by providing tentative implications that homosexuality fails at being a symptom of pathology. She suggests that homosexuality as a clinical entity might not exist, that homosexuality is within the normal psychological range as well as that "the role of particular forms of sexual desire and expression in personality structure and development may be less important than has frequently been assumed." (Hooker, 1957, p. 30). Homosexuality was thought to affect most if not all of the facets of the individual's personality, presumably because of the psychoanalytic influence in psychiatry which viewed homosexuality as a result of a developmental failure of the individual (Bayer, 1987). In her research, Hooker not only brought into question homosexuality as a clinical phenomenon but indicated that homosexuality was just a part, although an important part, of someone's personality and life which may or may not have consequences on the rest of person's psychological characteristics.

In light of her work, it became apparent that homosexuality was placed in diagnostic manuals on questionable scientific grounds. Unfortunately, the scientific evidence would fall on deaf ears were it not for the uprise of the gay rights movement since even after Kinsey's, Frank and Beach's and Hookers' studies indicated that homosexuality might not be pathological, some psychiatrist found ways to discredit their findings (Bayer, 1987). Regarding this, Bayer makes a crucial observation:

> The self-confidence with which those who viewed homosexuality as a pathology dismissed the findings of Kinsey and Ford and Beach as well as Hooker provides an important indication of the extent to which "facts" take on meaning only within the context of underlying conceptual schemes and do not in themselves have the capacity to compel fundamental changes in the way the world is viewed. (1987, p. 65)

As I have previously noted, similar to the scientific investigations in the beginning of study on homosexuality, the scientific data was interpreted to suit the already held views on the subject. Scientific data itself proved to be insufficient in changing the status quo of the psychiatric view of homosexuality. As Bayer concludes:

> As long as homosexuality was understood to be an "unrealistic adaptation" based upon "fear of heterosexuality", as long as it was seen as a deviation from the biological norm of the species, it was possible to characterize the data of critical investigators as scientifically irrelevant. Only when the basic perspective on the nature of normal sexuality began to change did that data begin to assume importance. When the change did occur, in large part because of the political struggle on the part of homosexuals, the evidence that had been available for more than two decades took on new meaning. (Bayer, 1987, p. 66)

These findings did, however, provide gay activists with strong arguments in the language psychiatrists themselves swore to speak in - the language of science. These studies significantly and impartially showed that there might be reasons to rethink the psychiatrists' stance on homosexuality. The rise of the homophile movement now had a strong ally in the form of scientific evidence that facilitated the atmosphere of change.

### ***1.1.2*** *Activist Work*

While crucial for facilitating the climate of change and gaining scientific allies such as Evelyn Hooker, the movement's politics of education, lobbying and litigation went only so far (Bayer, 1987). As Kameny proposed, the method of reasoning and logical argumentation "will not prevail, because most people operate not rationally but emotionally on questions of sex in general, and homosexuality in particular, just as

they do on racial questions." (*The Ladder: A Lesbian Review, May 1965, Vol. 9, No. 8 | Alexander Street Documents*, n.d.).

As Kameny urged to solidarization and protest, many complied and gay groups around US engaged in activities such as rallies, marches, picket lines and sit-ins, some of it inspired by the Civil Rights movement (Bayer, 1987)

The injustice, rage, and hurt from the mistreatment that has been building up for decades in homosexual circles finally culminated in a series of protests beginning on the 28th of June 1969 when during a police raid on the Stonewall bar in Greenwich Village the gays furiously, (and finally), fought back the police brutality and mistreatment. These became known as the 'Stonewall riots' which marked a new beginning of the gay rights movement, and which is considered by many to be the birthplace of gay rights as we know them today in many Western countries. Following the 'Stonewall riots', 'gay pride' demonstrations were organized. On the anniversary of the Stonewall riots and in many years to follow, 'gay pride' rallies take place with staggering numbers of participants. The Stonewall riots are celebrated and commemorated to this day on its anniversary by gay prides all over the world as a symbol of resistance and of the fight for acceptance.[2]

The anger and resistance of the gay community was not only taken to the streets but also to the meetings of the scientific communities, especially to the meetings of the American Psychiatric Association (APA). The APA published the official Diagnostic and Statistical Manual of Psychiatric Disorders (DSM) (American Psychiatric Association, 1980, 1994, 2013) which to this day serves as one of the main reference points in clinical and scientific practice. It contains every recognized mental disorder with their descriptions and lists of symptoms and is used extensively in diagnosis. The DSM in its first two iterations (DSM I and DSM II) had homosexuality listed as a mental

[2] June is usually considered to be the "Pride Month".

disorder. The gay activists saw changing that as one of the most important objectives in the fight for gay rights.

By 1973 the issue of whether homosexuality is a mental disorder and how it should be classified became a burning issue in the psychiatric ranks. The gay community advocates still applied pressure, but now as interlocutors in the dialogue by participating in Nomenclature committee meetings and bringing forth scientific evidence to justify their position, one notable appearance being that of Charles Silverstein (Bayer, 1987). In the May of 1973, some psychiatrists participated in the symposium entitled Should Homosexuality be in the APA Nomenclature? at the annual APA meeting which took place in Honolulu, Hawaii (Drescher, 2015). The symposium gained considerable attention from the psychiatric profession with participants in the thousands. The general atmosphere of the conference and the views discussed indicated that the homosexual dream of resolution from the mental disorder status would soon be a reality.

The Nomenclature Committee, a scientific body appointed by APA to tackle the classificatory issue of homosexuality was responsible for investigating and coming to conclusive results. Robert Spitzer was a chair of one of the subcommittees and he laid the scientific as well as conceptual groundwork for removing homosexuality from the DSM.

After careful consideration, the Nomenclature Committee decided homosexuality to be removed from the DSM II (Drescher, 2015). Many other fractions and committees of the APA followed suit and agreed with the decision respecting the authority of the APA appointed bodies (Drescher, 2015). In December of 1973, the APA's Board of Trustees accepted the recommendations and the decision of the Nomenclature Committee as they voted to remove homosexuality from the DSM II (Drescher, 2015). Finally, on the 15th of December they held a final vote with the Board and the removal of homosexuality won.

### *1.1.3* *The 'Sympathetic Ear' of Psychiatry*

Starting with the work of Evelyn Hooker, the part to emphasize for the sake of this thesis is not only her results but her methodology. Hooker was adamant to 'step into the shoes' of her homosexual subjects. She explored their world and what it means to be a homosexual male from their own point of view. She listened to the stories of homosexuals, their struggles, their problems, worries and suffering. She participated in and observed their lifestyles by attending gay gatherings, going to gay bars, and socializing with them. She was intent on finding out how homosexuality affects them, whether it is something that is problematic in itself or whether the problems are of the social nature. She was interested in their stories, in various aspects of their lives beyond their sexual proclivities. Unlike her predecessors, who approached the issue of homosexuality in white coats, she immersed herself in their world. Instead of assuming authority in telling them what was wrong with them, she listened. While her courage and tenacity were unparalleled being a straight woman in the 1950s exploring such a controversial subject in unconventional settings, her intellectual vigor and dedication to the truth were just as exceptional and admirable.

Thus, a lesson to be drawn from her work is that in approaching persons suffering from a condition it is important to study the patient's own position, to see how they are affected from their own point of view. That being said, I take it as a given that one has to keep to one's intellectual width, the background knowledge of the profession and one's role as a scientist, clinician or otherwise as well. I believe this kind of empathetic perspective is crucial in determining how the person is affected by a particular condition. Hooker's open-mindedness to listen to the experiences of homosexuals, to assume their point of view and her ability to empathize proved crucial not only for her own scientific contribution but for the significance her work assumed in the years to come with the rise of the gay rights movement.

Her work was of major importance to the gay rights movement as it showed that homosexuality did not "affect the psychological well-being of the individual" (p. 51) but that many of the non-clinical homosexuals were 'ordinary individuals' able to lead functioning lives (Bayer, 1987). The most important impediment to their functioning and to their potential psychopathology proved to be the stark social disapproval, exclusion, and stigmatization in the prevailing anti-homosexual climate.

This sensitivity to the individual's perspective would later become explicit in the form of the harm criterion. The story of depathologization of homosexuality, and Hooker's part in it strikingly illustrates the role and the importance of the harm criterion that will in years to come be introduced in psychiatry.

Another lesson to be inferred from the case of depathologizing homosexuality is the significance of conceptual work in psychiatry and the sympathetic ear of the clinicians and psychiatrists involved in patient care. With APA's announcement of removal of homosexuality from the list of mental disorders Washington Post came out with a headline titled 'Millions Cured'. Less medically surprising, but more philosophically important is the fact that millions were 'cured' by a conceptual rather than a medical intervention. Whether homosexuality was considered a disorder was a matter of the way the homosexuality was thought of and conceptualized and in turn how it fit into the definition of mental disorder. Homosexuality challenged the views and intuitions about the concept of mental disorder and urged for its reconsideration.

Homosexuality was viewed and defined as something inherently pathological, as the psychiatric profession had contact only with those homosexuals who had considerable comorbidities with other mental disorders. Based on this highly contestable and skewed view of homosexuality, it entered the diagnostic manuals and it fit the definition of mental disorder almost tautologically.

This was about to change with notable conceptual work of Robert Spitzer who proposed the reformulation of the notion of disorder in light of the homosexuality

debate. In the commentary entitled *The Diagnostic Status of Homosexuality in the DSM-III: A Reformulation of the Issue*, Robert Spitzer (1981) looked back on the events of the debate of 1973 to provide the theoretical rationale for the exclusion of homosexuality from the DSM II, as well as explaining the changes decided on in the subsequent edition, the DSM III, and the process leading up to them.

In discussing the definition of mental disorder, he recognized that most, if not all, disorders carry with them subjective distress or "were associated with generalized impairment in social effectiveness or functioning." (Spitzer, 1981, p. 211). In other words, most conditions involve either disturbing and unfavorable psychological states such pain, anguish, anxiety and so on, or they are characterized by inability to function in various important areas of live. This became known as the harm criterion in mental disorders which assumed its standard formulation of 'distress or disability' in the DSM III:

> In DSM-III each of the mental disorders is conceptualized as a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is typically associated with either a painful symptom (distress) or impairment in one or more important areas of functioning (disability). In addition, there is an inference that there is a behavioral, psychological, or biological dysfunction, and that the disturbance is not only in the relationship between the individual and society. (When the disturbance is limited to a conflict between an individual and society, this may represent social deviance, which may or may not be commendable, but is not by itself a mental disorder.)(American Psychiatric Association, 1980, p. 6)

Thus, disorders included in the DSM would have to satisfy either of the two horns of the so-called harm criterion. Psychiatrists would have to determine whether a particular condition carries with them subjective distress in the patient or whether

the patient is impaired in important areas of functioning (disability). Uncovering harm as the common denominator of mental disorders gave Spitzer a conceptual backdrop against which to exclude homosexuality from the notion of mental disorders. This both initiated and facilitated a paradigm shift in psychiatry.

With regards to homosexuality, Spitzer emphasizes that distress is not inherent to the condition (Spitzer, 1981). Considering homosexuality in comparison with anxiety where subjective distress is inherent and arguably the most salient symptom, in homosexuality there are none such symptoms. Granted, homosexuals often suffer distress which is brought about mainly by external factors such as disapproval, discrimination, abuse or threat of abuse and so on, as well as "auto-homophobia" which represents the internalized negative social view on homosexuality.

The other element, referred to as disability, states that a condition is associated with "impairment in one or more important areas of functioning". Here, Spitzer makes several interesting remarks. First, to consider homosexuality as failure in an important area of functioning would be to consider heterosexuality a norm from which homosexuality deviates, which is a claim strongly contested by the gay community (Spitzer, 1981).

Furthermore, Spitzer keenly notices an important feature in the definition of mental disorder. Whether we take heterosexual functioning as a norm is a value judgment[3] (Spitzer, 1981). The same goes for determining which areas of functioning are 'important' claiming that there is always a value judgment in adjudicating whether we consider something important and in what way. While many experts agree sexuality to be one of the important areas of functioning, the bone of contention lies

[3] Value judgments can in this context be contrasted with scientific, i.e. value-free judgments. Pretheoretically, the notion of mental disorder produced by the American Psychiatric Association was generally taken to be a beacon of scientific authority having value-free pretensions.

in whether heterosexuality is the 'important' kind, whether heterosexuality is the norm (Spitzer, 1981). Spitzer (1981) proposes that the concept of 'inherent disadvantage' could be used in determining whether some area or property of functioning is important:

> For example, an individual who is not able to test reality because of delusions or hallucinations, or an individual who is not able to function occupationally because of depression, is at a clear inherent disadvantage in efforts to satisfy basic biological and psychological needs. (Spitzer, 1981, p. 212)

Conditions that are in themselves, intrinsically, disadvantageous in everyday functioning of the person and in fulfilling their basic needs are those that are harmful. Again, homosexuality may lead to impairment in important areas of functioning, and thus be harmful, but as a result of discriminatory social climate. Thus, homosexuals are not harmed by their condition, in virtue of their condition. And, since they are not harmed by their condition, they are not mentally disordered.

The argument to exclude homosexuality from the list of disorders might be reformulated to look like this:

> 1. Mental disorders are "typically associated with either a painful symptom (distress) or impairment in one or more important areas of functioning (disability)." (Spitzer, 1981) (The harm criterion)
> 2. Homosexuality does not cause distress or impairment in social effectiveness or functioning since many homosexuals are capable of living healthy and fulfilled lives. Homosexuality is not harmful in itself.
> 3. Therefore, homosexuality is not a mental disorder.

Extracting the harm criterion which seemed to have been in the background of many disorders, and making it salient, led to a shift in thinking about mental disorders in psychiatry. It seemed to have embodied the perspective I have lauded throughout this chapter, namely, the perspective of the patient. The harm criterion is the realization, the manifestation of the individual's perspective and assessment of their life from their own point of view. It is the perspective which has been pointed at by the scientific research on homosexuality, fought for by the gay activists and, in time, recognized by the psychiatrists of the APA.

Spitzer in his conceptual work on the notion of disorder insightfully reflects on and makes highly significant distinctions, the harm criterion being just one of these. Specifically, in his 1981 paper he starts off discussing homosexuality by saying that the concept of disorder is man-made, designed to be of some use to us maintaining that the nature and the function of the concept of disorder is to helps us identify and treat people who are in need of some help (Spitzer, 1981). Furthermore, he claims that in determining whether someone lacks in important areas of functioning, which also has implications on whether we consider someone as disordered, does not rest on facts alone but instead leans heavily on values (Spitzer, 1981). Therefore, according to Spitzer the concept of disorder is value-laden and cannot be adjudicated by simple fact-checking.

As we shall see in the next chapter, these two claims are pertaining to particular positions in the debate that are central to the philosophy of psychiatry. This is the discussion on the definition of mental disorders. This debate tackles the questions of how disorders should be defined, what should their scope be, who they should encompass and why, what elements or criteria should the notion of disorder contain and so on.

## 1.2 Harm in Diagnostic Manuals

With the changes introduced in the DSM – III including harm as distress or disability in the definition of mental disorder, harm was accepted and featured in the subsequent edition of the manual as well - the DSM – IV. The concept of mental disorder in DSM – IV, much like in the previous edition was defined as such:

> In DSM-IV, each of the mental disorders is conceptualized as a clinically significant behavioral or psychological syndrome or pattern that occurs in an individual and that is associated with present distress (e.g., a painful symptom) or disability (i.e., impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom. (American Psychiatric Association, 1994, p. xxi)

However, even though DSM – IV featured a definition of mental disorder they put a caveat in place. The APA emphasized that although DSM "provides a classification of mental disorders, it must be admitted that no definition adequately specifies precise boundaries for the concept of "mental disorder." (American Psychiatric Association, 1994, p. xxi). Rather, "The concept of mental disorder, like many other concepts in medicine and science, lacks a consistent operational definition that covers all situations." (American Psychiatric Association, 1994, p. xxi). Thus, the definition should have been viewed as a suggestive guideline for thinking about the notion of mental disorder. Thus, even though the APA offers a definition they take the stance of laudable intellectual humility. The definition offered in the DSMs is not the be-all-and-end-all when it comes to defining mental disorder. Rather, it is the closest or the most accurate approximation that they have come up with.

Another prominent diagnostic manual, International Classification of Disease (ICD) developed by the World Health Organization (WHO), in its 10th edition offers a similar definition of mental disorder to that of the DSM. To them, mental disorders are “clinically recognizable set of symptoms or behaviour associated in most cases with distress and with interference with personal functions.” (World Health Organization, 1992, p. 5). Consistency between manuals was something the both groups aimed for as “Those preparing ICD-10 and DSM-IV have worked closely to coordinate their efforts, resulting in much mutual influence.” (American Psychiatric Association, 1994, p. xxi).

However, the romance of harm as distress or disability being a necessary criterion of mental disorders has been short lived. Such view on harm faces significant difficulties on both prongs. Looking at distress, it has been pointed out that the notion is vague and underspecified (Amoretti & Lalumera, 2019, p. 4). As Amoretti and Lalumera (2019, p. 4) indicate, even though distress is present and often evoked in the diagnostic manuals, it is “defined neither in DSM-5 nor in the most recent revision of the International Classification of Diseases (ICD-11).”. Phillips as well indicates that even though both diagnostic manuals, the DSM as well as the ICD, employ the notion of distress “neither the DSM-IV nor the ICD-10 provides a definition of the term, so there can be a wide range of interpretations of the corresponding diagnostic criteria.” (Phillips, 2009, p. 92).

There is yet another issue with distress. Distress is an expected response to life stressors, for example, loss of a loved one, dangerous or life-threatening situations. Besides this nonpathological distress, there is maladaptive, pathological kind of distress. As Phillips (2009, p. 91) writes, there are a couple of ways authors explain distress, and for the purposes of this work I will take just one (for discussion see Phillips (2009)). This one argues that distress is a “transient phenomenon related to specific stressors that subsides when the stressor disappears or as the individual

adapts to the stressor." (Phillips, 2009, p. 91). Now, if distress persists or is inappropriate in severity or longevity after the termination of the stressor, then we start thinking about it in terms of a pathological symptom in the purview of mental disorder (Phillips, 2009, p. 91). For example, the distress (manifested as intense fear, nervousness, agitation, shallow breathing, sweating) of a fireman running into a burning building is an appropriate, adaptive response to the situation. However, if that same fireman months after an event feels agitation, 'jumpiness', uneasiness, nervousness and intense fear just by sitting at home watching TV, or the response is triggered by something like the sound of sirens, then we can start talking about a pathological condition developing.

The fireman case serves merely as an illustration aimed to show that distress is not inherently harmful but that there are healthy variations of it. However, acknowledging that leads us to the problem of how to differentiate between these pathological and nonpathological instances of distress. In the reality of the medical practice, things are not nearly as simple and discernible as the fireman case. In fact, we do not possess a good way for distinguishing between maladaptive and adaptive distress: "the diagnostic systems do not assess the degree of distress and do not provide further clarification about the cut-off between distress that is and is not diagnostically important." (Phillips, 2009, p. 92).

The second prong of harm, the disability criterion, is problematic as well. The discussion on disability in philosophy of medicine is a longstanding and elaborate warranting the attention and engagement that is beyond the scope of this work. There is, however, a well-known dichotomy in the literature on disability that we can address, between medical and the social model of disability (Wasserman & Aas, 2022). The medical model of disability "explains disability disadvantage in terms of pathological states of the body and mind themselves. It regards the limitations faced by people with disabilities as resulting primarily from their bodily differences."

(Wasserman & Aas, 2022, sec. 1.1 Medical and Social Models). But the medical model of disability is "rarely defended explicitly, but aspects of it are often adopted unreflectively, when health care professionals, bioethicists, and philosophers ignore or underestimate the contribution of social and other environmental factors to the limitations faced by people with disabilities." (Wasserman & Aas, 2022, sec. 1.1 Medical and Social Models). The medical model in the disability literature is considered as suboptimal and a bit outdated and as it says earlier, rarely adopted in discussions that address the issue of disability reflectively and explicitly. Conversely, the social model of disability "explains the characteristic features of disability in terms of a relation between an individual and her social environment: the exclusion of people with certain physical and mental characteristics, or "impairments", from major domains of social life." (Wasserman & Aas, 2022, p. 1.1 Medical and Social Models). The social model of disability is the one favored in the literature. It is also the one adopted by the prominent manuals such as the *International Classification of Impairments, Disabilities, and Handicaps* which defines disability as *"*any restriction or lack (resulting from an impairment) of ability to perform an activity in the manner or within the range considered normal for a human being" (World Health Organization, 1980, p. 28) thus putting emphasis on behavior and performance of a person in relation to its environment rather than on their bodily or mental configurations. The social model is also the one adopted by Amoretti and Lalumera in their analysis of harm in psychiatry where they argue that disability is "typically considered to be an intrinsically relational concept, which involves an environmental and social component" (Amoretti & Lalumera, 2019, p. 5). Accordingly, they point out that disability is viewed as a result of a disorder rather than being a disorder in itself (Amoretti & Lalumera, 2019) (see also World Health Organization (1992)). I take these as sufficient reasons for adopting the social model of disability for the purposes of harm in psychiatry. Disability is thus a dynamic criterion that negotiates between the

person and the environment. It is not intrinsic to the patient but is "intrinsically relational" depending on the way the person interacts with the environment. As such, the criterion of disability is problematic in the process of diagnosis. It seems that disability manifests itself in response to a condition, as a result of one, rather than as an integral component of a mental disorder. This means that the notion of disability includes and entails problems outside of the scope of diagnostic criteria embedded in diagnostic manuals and it makes it that much harder to properly assess and diagnose mental disorders. It seems to carry with it more conceptual burden than it resolves.

Furthermore, the problems with harm criterion do not end there. Besides the lackluster and underdefined notions of distress or disability, there has been a shift in the latest edition of DSM (2013) where harm went from a necessary to a usual criterion of mental disorder. Here is the definition in the DSM – 5:

> A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviant behavior (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above. (American Psychiatric Association, 2013, p. 20)

We see the shift from mental disorders as being “associated with present distress or disability” to them being “usually associated with significant distress or disability”. It now means that harm does not have to be present for diagnosing a person with a mental disorder.

While at first glance the change in the definition seems more of as a technicality, authors like Cooper (2020) recognize the significance of such a move and deem it as reflective of a change regarding the meaning of harm in psychiatry. Cooper says that “For about forty years, from the early 1970s onwards, there was a broad consensus amongst mental health researchers and clinicians that harm is essential for a condition to be a mental disorder.” (R. V. Cooper, 2020, p. 145). However, she emphasizes that the consensus has broken down. She describes the way the change has transpired. She claims that there were two competing definitions of mental disorder under consideration in the making of DSM - 5. One was an “iteration of the DSM previous definition” that took harm as a necessary condition of mental disorder (R. V. Cooper, 2020, p. 146).

The other definition under consideration aimed at making the definition of mental disorder compatible with the literature on “physical disability employed by the World Health Organization” (R. V. Cooper, 2020, p. 146). As such it aspired to make the definition of mental disorder value-free in the form of biological dysfunction, in order to be compatible with the distinction between ‘impairment’ and ‘disability’ in the disability literature (see Cooper (2020) for discussion). As a compromise between these two options of defining disorder, the definition of disorder that does not necessitate harm entered the DSM. In light of this, Cooper concludes that harm in psychiatry does not represent what it used to, going from a necessary to a usual criterion of disorder. In addition, rather than distinction being a dead letter, Cooper notices that it is evocative of the way clinicians talk (and presumably think) about specific cases of mental disorders. This change is noticeable

in the way tic disorders are treated. As Cooper writes, according to DSM IV (American Psychiatric Association, 1994), to be diagnosed with a tic disorder one would have to be harmed by it. However, in DSM - 5 (American Psychiatric Association, 2013) one does not have to be harmed by tics for them to be considered a disorder. She concludes: "The psychiatrists are now happy to claim that (i) tic disorders often cause no harm, and (ii) tic disorders are disorders." (R. V. Cooper, 2020, pp. 147–148). Therefore, Cooper points out that harm has been losing ground in psychiatry.

As we have seen, harm in diagnostic manuals is problematic in many ways. Both elements of harm, distress and disability face considerable challenges. Distress is underspecified and vague with none of the manuals offering a clear and in-depth definition. Furthermore, there is no conclusive way to differentiate between pathological and nonpathological distress. Concerning the disability element, it is considered to be intrinsically relational meaning that in determining whether someone is harmed depends on the outside factors and the person's relation to them which seems counterintuitive in the light of psychiatric diagnostics.

Besides diagnostic manuals, harm has been featured in the discussion on the definition of mental disorder in philosophy of psychiatry.

### 1.3 Harm in the Debate on Mental Disorders

Around the same time as the changes in mainstream psychiatry were starting to take place – depathologization of homosexuality, changes to the definition of mental disorder in diagnostic manuals - analytic philosophy of psychiatry was becoming a prominent field of philosophical inquiry. Philosophy of psychiatry can be defined as a "familiar, and expanding, field of inquiry that often takes the tools of contemporary epistemology, ethics, or philosophy of mind, and applies those tools to the subject matter of psychiatry—psychopathology, mental health services and research, and

mental health law and policy." (Moseley & Gala, 2015, p. 2). Basically, it is using philosophy to investigate the field of psychiatry.

Now, the reasons behind the revival of analytic philosophy of psychiatry purportedly had to do with another movement of the 1960s – the anti-psychiatry movement (Thornton, 2007, p. 11)[4]. This was the time when the anti-psychiatry movement was gaining prominence throughout the most of Europe and Anglo-American countries, highly impactful in countries such as Italy, Netherlands, and the UK.

Anti-psychiatry is "a broad-based movement that questioned the legitimacy of standard psychiatric theory and practice." (Whitley, 2012, para. 1). Specifically, the movement questioned the legitimacy of psychiatric categories, diagnoses and all in all questioned the idea of mental disorder. Several notable anti-psychiatry works and authors of 1960s "set the stage" for the movement such as "Ronald Laing's *The Divided Self*, Erving Goffman's *Asylums*, and Thomas Szasz's *The Myth of Mental Illness"* (Grob, 2011, p. 401). Szasz's The Myth of Mental Illness epitomized the movement.

Szasz (1960) claimed that mental illness is a myth, propagated by the psychiatric authority for the purposes of social control. He drew a distinction between psychiatry and bodily medicine arguing that there is no such thing as mental illness – the phenomenon was either a brain disorder or it was a problem of living. If it were the former, then it would have been approached as any other form of brain disease – surgical, pharmacological or some other kind of intervention directed at brain lesions or similar conditions treated by neurology. If it were the latter, the problem of living, then it would not fall under the purview of medicine and psychiatry but would be a social condition. He thought that it is no place of medicine and of psychiatry to treat

[4] Thornton postulates that analytic philosophy of psychiatry "has been reborn in large part as a response to the rise of anti-psychiatry movement in the 1960s"(2007, p. 11).

social issues and personal problems, but that medical intervention in such cases represents a repressive misuse of medical authority. According to Szasz, psychiatry was "little more than a pseudoscience" lacking reliability and validity but which "embodied a set of value judgments that imposed a particular view of bourgeois reality (with all of its vested interests) on a minority" (Grob, 2011, p. 403). Thus, psychiatry "was merely an instrument of social control" and as such it "constituted a threat to individual liberty in a free society" (Grob, 2011, p. 403).

The crux of the issue between psychiatrists and anti-psychiatrists was the concept of mental disorder. As Thornton writes: "the anti-psychiatry view that mental illness does not exist has immediate repercussions for the justification of mental health care practice.", making the analysis of mental illness not only one segment of interest in philosophy of psychiatry but its "key question" (Thornton, 2007, pp. 11–12). This is indeed what has been the burning issue in philosophy of psychiatry for the last forty years – what are mental disorders and how to define them?

This issue expanded into a longstanding and heated debate about mental disorders in philosophy of psychiatry which I will address in the following chapter. Especially pertinent for this work is that harm has been one of the elements featured in the debate on mental disorders and many authors include harm in their analysis (Bortolotti, 2020; R. V. Cooper, 2002; Glover, 1970; Reznek, 1987; Wakefield, 1992).

One of the earlies accounts of harm in psychiatry can be found in Laurie Reznek's *The Nature of Disease* (1987). Reznek (1987, p. 134) starts his analysis by proclaiming that "the most plausible Normativist Theory defines diseases in terms of harm", however, he goes on to say that in order to investigate the notion of mental disorder in terms of harm, we must assess whether harm is an evaluative notion. From the perspective of the debate on mental disorders this seems a bit strange as harm has been taken to be, almost tautologically, a normative element that is value-laden. Reznek discusses whether harm is naturalist or normativist but finally arrives to

treating harm as a normative notion. I take this to be the starting point of laying out Reznek's position.

On the Normative Theory of harm "our good or welfare is constructed from ideal-regarding interests. We have an interest in just those things that we judge will enable us to lead a good or worthwhile life" (Reznek, 1987, p. 150). The concept of ideal-regarding is traditionally contrasted with the concept of want-regarding. The distinction was introduced by Brian Berry who defined want-regarding principles as those that take into account everything a particular person wants whereas the idea-regarding principle puts the wants of a person in perspective of their overall welfare or well-being (McLean & McMillan, 2009, sec. ideal–regarding principle). For example, smoking might be a want-regarding interest because it offers momentary pleasure, but it is not an ideal-regarding interest since it has been linked to development of serious respiratory illnesses. When it comes to Reznek the dichotomy between want-regarding and ideal-regarding refers to the theories of interest with his normative account favoring the ideal-regarding theory. According to Reznek (1987, p. 150), "someone is worse off, if and only if, his interests are impaired" and someone's interests are impaired "if and only if, his leading a good or worthwhile life is interfered with". Combining these two propositions we get to the final formulation of Reznek's (1987, p. 153) account of harm:

> X does A some harm if and only if X makes A less able to lead a good or worthwhile life

Reznek takes this account of harm as a main component of the notion of mental disorder. He incorporates it into a definition of mental disorder he arrives at after an extensive analysis:

> A has a pathological condition C if and only if C is an abnormal bodily/mental condition which requires medical intervention and which harms standard members of A's species in standard circumstances (Reznek, 1987, pp. 163–164)

I will not go into Reznek's definition of mental disorder at length but will focus my attention to the notion of harm. The account of harm on his view is a bit problematic because it rests on the notion of good or worthwhile life. This is conceptually burdensome since the notion of worthwhile life is controversial. There is a multitude of various understandings of what a worthwhile life entails, both in philosophy and in everyday life. In contemporary democratic societies we are accustomed to personal freedoms that allow for an array of different visions of the good life. This makes the concept vague and problematic to define as any or most of substantive theories of a good life might infringe on some personal freedoms or be contrary to some visions of the good life. If saying that friendship and affiliation makes life good and worthwhile, then the person who chooses to live life as a recluse in pursuit of spiritual awakening does not live a good life, or even their vision of the good life harms them. This seems wrong. Who are we to have the right to say which lives are worthwhile? How are we to judge this without also impinging on basic human liberties and freedoms to live as they want? Thus, such views that propose substantive theories of what it means to live a worthwhile life are vulnerable to the objection of paternalism. Paternalism is "the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm." (Dworkin, 2020, para. 1). Paternalism is seen as negative, oppressive, and stifling to the individuals in a free and democratic society. It is what we as a society try to avoid.

Cognizant of these problems with the concept of worthwhile life, Reznek offers a few caveats. He argues that "there is not just one sort of worthwhile life – a worthwhile life might consist in people realizing their different worthwhile potentials" (Reznek, 1987, p. 151). Thus, for one person that might entail living a life of quiet contemplation while for another it might mean engaging in adventurous and exciting experiences. Therefore, Reznek accepts that the same thing that would make one person better off can make another worse off.

Additionally, he argues that "it is difficult to think that a life would be worthwhile unless one had positive desire for it." (Reznek, 1987, p. 151). Positive desire entails that the life person leads is the one they affirm subjectively. In other words, the good or worthwhile life is the one person desires for its own sake. For example, someone might think that success and career-driven life is the one worth for its own sake but if the person does not desire or want such life and sees no value in it then our intuitions detect a problem. We see disparity between the life they "ought to" live versus the life they "want to" live. As Reznek (1987, p. 151) says, "most of us think that in order for a life to be worthwhile, it must be desired for its own sake". Therefore, personal affinities and desires intuitively hold more weight than theoretical implications of what one ought to do or desire.

Furthermore, Reznek argues that even though the good life cannot be equated with just satisfaction of desires or the achievement of happiness, worthwhile life seems to entail at least some of these things: "A person's good or welfare consists in the satisfaction of worthwhile desires and the enjoyment of worthwhile pleasures" (Reznek, 1987, p. 151). While living a worthwhile life is not all about desire-satisfaction or pleasure alone, intuitively, lives that are completely devoid of such experiences do not seem to line up with our intuitions on what is a good life.

Thus, these caveats weaken the paternalism objection by stating that there are various kinds of lives that are worthwhile, that the person should have the positive

desire, that they want to live the life that is judged as worthwhile, and that it does include a certain amount of pleasure and satisfaction. This makes the account more palatable as it takes into account person's actual preferences, desires, and allows for a range of 'good lives'. Unfortunately, Reznek is still not able to evade the problem of defining the worthwhile life and the question of what such life would entail. These questions remain unanswered. His account ties harm to a theory or theories of well-being, however, Reznek does not offer further elaboration by subscribing to any of the theory of well-being. Thus, Reznek's account leaves something to be desired when it comes to harm.

Another author that can be considered as one of the early mentions of harm is Jonathan Glover's and his book Responsibility (1970). In it he argues that "we should only regard someone as ill, and hence in need of treatment, where he is in a physical or a mental condition that is harmful to him" (Glover, 1970, p. 119). He emphasizes that the harm condition "sets limits to what can be treated as mental illness" (Glover, 1970, p. 120) meaning that then a condition has to be harmful to be a mental disorder rather than a mere "deviance" from normality. It is a necessary line of defense against the psychiatric mistreatment and misdiagnosis for which the history of psychiatry is infamous. Not all things that are harmful are mental conditions, but harm is a necessary, yet not a sufficient condition of mental disorder, according to Glover (1970, p. 120). So, how does Glover then view harm?

As a starting point, Glover takes Bentham's analysis of pain which divulges into a taxonomy of twelve kinds of pain. Glover (1970, p. 121) concludes that there is a "prima facie case for saying of a condition that it is harmful to someone whenever it is seen as unpleasant or unwanted by him, or whenever it deprives him of pleasure or of the satisfaction of his wants". Granted, there may be cases where the harm or deprivation of benefit can have a "compensatory benefit" or they "lead to avoidance of some greater harm" (Glover, 1970, p. 122), for example, avoiding pleasure of

sugary, salty and fatty foods to prevent future harm of cardiovascular diseases. However, unpleasant or unwanted conditions, writes Glover, are always prima facie harmful and only upon citing some further reasons of greater benefit or harm avoidance can the harm and benefit be levelled out.

Glover emphasizes a very important point in his analysis of harm. He postulates that it is "want-regarding" rather than "ideal-regarding" – "to be harmed is to be given what is unpleasant or unwanted without compensating benefit; it is not to be deprived of what others think is 'good for one' or to be given what they think is 'bad for one'" (Glover, 1970, p. 122). He provides example of this:

> Thus, you may disapprove of my drinking or gambling, but your view that one is a better sort of person if one does not drink or gamble is not sufficient to show that I am harmed by these habits. To establish this, you would have to show that drinking, gambling or their consequences involve reductions in my pleasure or want-satisfaction. (Glover, 1970, p. 122)

Recognize that considering harm want-regarding is different from Reznek's account where he supports the ideal-regarding view of harm. I would be keen to side with the latter as it seems to be one of the most important characteristics of harm, one that I have underlined throughout this work. Things started to change for the condition of homosexuality only when their perspective was thoroughly and extensively explored. Harm takes into account the first-person perspective of the individual who has a particular condition. It considers what their life looks like from their own point of view and how the condition factors into their life.

However, there are problematic aspects of adopting this view of harm as a relevant psychiatric criterion. Glover (1970, p. 122) in his work recognizes, what he calls, the problem of 'unconscious deprivations'. These are circumstances in which a

person may be suffering real deprivation but is not aware of it. In other words, their life might be marked by real deprivation but because they do not know of any other way they may just be used to their condition and apathetic towards the prospect of changing it. For example, think of a person that suffers from social anxiety. The severity and the prolonged effect of her condition may have steered her life in various directions, possibly contrary to her real wants or desires. She may have chosen a job where the human contact is minimal, she may have developed hobbies and interests that do not include participation of others, her social life might have been replaced by to the virtual world or she may be used to being without human contact and so on. This leads Glover to conclude that we are not always the final authority on whether we are harmed. However, he condemns the extreme alternative view "by which one assumes that men must be unconsciously harmed or deprived whenever they have a form of life that one would not oneself choose" (Glover, 1970, p. 123). Thus, in assessing harm we should come to a compromise between the 'first person perspective', the want-regarding perspective of the individual and the 'unconscious deprivation' they may be the victim of. Thus, "in a state of prolonged withdrawal", as I have exemplified with the case of social anxiety, "the condition is harmful" if the person seems distressed (if they are experiencing unpleasant of unwanted states) or if there is strong evidence of unconscious deprivation (Glover, 1970). Determining the former, whether the person is distressed is problematic, but not as much as determining unconscious deprivations. This is where the situation gets significantly more problematic.

Glover offers a clue for resolving it. In figuring out whether a person is suffering from such deprivation: "The best evidence can only be provided by those who have experienced both normal life and the state of withdrawal" (Glover, 1970, p. 124). In this context, "it is only because they say they are glad of the change that we have any right to describe it as a 'cure'" (Glover, 1970, p. 124). Thus, an option for the

cases of unconscious deprivation is making the person aware of the options that the resolution of a state might afford them or maybe the testimony of other people in these situations whose experience might show that they are better off after treating a condition. Regardless of these suggestions, unconscious deprivation remains a serious problem for the notion of harm in psychiatry in general.

Besides the problem of resolving unconscious deprivation, Glover proposes guiding principles for determining which conditions are harmful. Thus, assuming the position that "only those bodily or mental conditions that harm the person who has them" (Glover, 1970, p. 124) count as mental disorders,

> the condition must be 'directly' harmful: the harm must result from my mental or physical state itself, without the reactions of other people to my state being a necessary condition of harm. (…) Physical or mental suffering are harmful in themselves, even if they have no further harmful consequences. (Glover, 1970, p. 124)

The second condition that Glover puts forth has to do with the degree and the duration of the condition. According to him, "a day or so of mild depression on failing an examination might seem too insignificant to call for treatment of mental illness" while the more prolonged and severe cases might warrant psychiatric treatment. Glover (1970, p. 125) concludes that recognizing mental disorders as being "without a sharp boundary" may bring mental disorders closer to nonpathological conditions, them being a matter of degree rather than the matter of kind.

Glover brings out important points into our analysis of harm. He introduces the importance of the first-person perspective and highlights an important challenge for harm in the form of 'unconscious deprivation'. He proposes a way of resolving unconscious deprivation from the perspective of the patient by stating that we can

talk of ‘cure’ when the patient evaluates their condition as pathological and in need of assistance after the fact of treating a condition. In other words, when the person obtains a point of comparison between the problematic condition and their life devoid of this condition. In this Glover sets a good foundation for thinking about harm. However, I believe his account as well lacks a bit of elaboration. How do we determine the relation between unpleasant or unwanted states and their compensating benefit? What if a person is in a state of unconscious deprivation but nonetheless experiences pleasure and wanted experiences?

Besides Reznek and Glover, several authors later on introduce harm in their analysis of mental disorder. One such author is Jerome Wakefield who argues for the harmful dysfunction account of mental disorder:

> A condition is a disorder if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some internal mechanism to perform its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism (the explanatory criterion). (Wakefield, 1992, p. 384)

Here I focus on the harm component of his account. He argues that dysfunction is not enough to consider some condition a disorder but that “the dysfunction must also cause significant harm to the person under present environmental circumstances and according to present cultural standards.” (Wakefield, 1992, p. 383). He provides an example of dysfunction in one kidney which is not considered to be a disorder because it does not have an effect on “the overall well-being of a person” and physicians can remove a kidney from a live donor without

causing them harm (Wakefield, 1992, p. 384). He goes on to say that "The requirement that there be harm also accounts for why albinism, reversal of heart position, and fused toes are not considered disorders even though each results from a breakdown in the way some mechanism is designed to function." (Wakefield, 1992, p. 384). Thus, even though according to Wakefield disorders must involve a biological dysfunction, not all biological dysfunctions are considered to be disorders. Only those dysfunctions that cause harm are disorders.

Wakefield's view of dysfunction is a breakdown in some evolutionary selected mechanism. He mentions two reasons for the relation of naturally selected mechanisms and harm. The first reason is that our naturally selected mechanisms and traits were those that might have benefitted us in our evolutionary past but it may be the case that they do not benefit us anymore in our contemporary way of living. For example, the phenomenon of premature ejaculation (PE) might have been evolutionary beneficial for proto hominids:

> because of their great speed, they were able to complete their copulations before the females and the other males could effectively deter their actions. With their ability to copulate quickly, they also had more opportunities to have sex with a greater number of females within a given period of time (an important point if the female protohominids had estrous periodicity). The fast ones, therefore, had more partners and more opportunities to impregnate, and, thus, were able to pass on their propensity for speedy ejaculation (perhaps a penis with a highly sensitive corona or frenum) to a greater percentage of the next generation. (Hong, 1984, p. 116)

However, even though premature ejaculation might have been evolutionary beneficial, in contemporary times it is seen as a downside rather than a virtue. It is included in the diagnostic manuals such as DSM as “Premature (Early) Ejaculation” (American Psychiatric Association, 2013, pp. 443–446). One of the defining criteria is the distress/bother, i.e. harm the phenomenon causes to the person (Rowland & Cooper, 2022, p. 2). As Rowland and Cooper (2022, p. 2) note, bother or distress has been used to encompass “all ‘negative consequences of PE’” including psychological, behavioral and interpersonal effects that the phenomenon has on the person and their partner. These include:

> descriptors such as bothered, dissatisfied, anxious, concerned, depressed, frustrated, ashamed, self-disgusted, embarrassed, and others. Negative behaviors have included, but not been limited to, verbal catastrophizing about the negative impact of PE, avoidance of physical intimacy, or profuse postcoital apologizing. (Rowland & Cooper, 2022, p. 2) (see also Althof, Mcmahon & Rowland (2022))

Thus, even though premature ejaculation might have been beneficial once, it is now causing significant psychological and interpersonal problems and as such is regarded as a disorder.

The second reason Wakefield connects the dysfunction requirement and harm has to do with the fact that “natural selection of a mechanism occurs when organisms that possess the mechanism have greater reproductive fitness than organisms that do not possess the mechanism” (Wakefield, 1992, p. 384). However, small decreases in reproductive fitness may occur which from an evolutionary perspective might seem important but they might not be harmful when it comes to disorder. Take for example the ability to postpone ejaculation and have significant control over it. This might

seem from the evolutionary perspective as a setback in reproductive fitness but in contemporary setting it might be actually beneficial to the person as it might be correlated with relationship satisfaction of both partners and their mutual bond, positive self-image and self-esteem and so on. Thus, Wakefield concludes that "The mental health theoretician is interested in the functions that people care about and need within the current social environment, not those that are interesting merely on evolutionary theoretical grounds." (1992, p. 384).

Wakefield emphasizes important points about the social impacts and preconditions in assessing biology and fitness. Even though he holds the notion of biological dysfunction in high esteem when it comes to mental disorder, he recognizes that the evolutionary explanation of conditions is not the be-all-and-end-all when it comes to the notion of disorder. Instead, mental disorders are only dysfunctions that are "socially disvalued" (Wakefield, 1992, p. 384) meaning that only dysfunctions that result in harm are mental disorders.

There are at least three issues with Wakefield's view of harm. He himself alludes to the first one:

> Note that in this article I have explored the value element in disorder less thoroughly than the factual element. This is in part because the factual component poses more of a problem for inferences about disorder and in part because the nature of values is such a complex topic in its own right that it requires separate consideration. (Wakefield, 1992, p. 384)

He does not dedicate as much attention to harm as he does to biological dysfunction in the original paper from 1992. However, he has been defending his theory of mental disorder as harmful dysfunction for the last thirty years and he has not yet addressed and elaborated on the issue of harm (at least not to my knowledge).

Therefore, his theory of harm remains underdeveloped and while he ardently defends the harm component of his account, as seen in Wakefield and Conrad (2019, 2020), his analysis of harm is still lacking.

Another problem with Wakefield's view of harm is its connection with biological dysfunction. He considers mental disorders to be conditions where biological dysfunction *causes* harm. Only in these cases can person be diagnosed with a disorder. However, tying harm to biological dysfunction is a bit problematic. Mental disorders are often complex phenomena meaning there can be several mechanisms at play resulting in different manifestation of symptoms. There is no guarantee that failure of one mechanism is the one that is biological dysfunctional and then causes harm. For example, major depressive disorder (MDD) is such a phenomenon that its pathophysiology is not yet completely clear. When it comes to the major depressive disorder:

> Research has implicated several mechanisms including altered serotonergic, noradrenergic, dopaminergic, and glutamatergic systems, increased inflammation, HPA axis abnormalities, vascular changes, and decreased neurogenesis and neuroplasticity. (Dean & Keshavan, 2017, para. 2)

These mechanisms which are responsible for the major depressive disorder might work together and be interrelated belonging to the same condition but they also might be independent of each other and consists of several malfunctioning mechanisms. Dean and Keshavan (2017, para. 4) conceptualize depression "as resulting from an interactive matrix of pathophysiological mechanisms" where the "alteration in the matrix at any node has the ability to trigger the entire cascade of biological effects.". This makes it hard to pinpoint where exactly biological

dysfunction lies, what it consists in and even harder to determine which part results in harm. After all, what we think of as depression now might not even be a unitary phenomenon but "depressive disorders may represent several subsyndromes or even numerous discrete disorders." (Dean & Keshavan, 2017, para. 4). The point is that we are still not sure and do not know the underlying mechanisms and its workings when it comes to the major depressive disorder. Dysfunction may even be an emergent property supervenient on the workings of numerous mechanisms. Regardless, it seems we can be privy to whether the person is harmed by the condition, and we may even be able to treat it pharmacologically, even if through a trial-and-error process of prescribing medication, as sometimes that may be the best that we have at our disposal.

Lastly, the problem of Wakefield's view of harm is that a condition is harmful if it is negatively evaluated by a person's culture. Thus, only those conditions that are socially disvalued are harmful and social standards play a major role in his view of harm. This is worrisome as once social standards were such that they deemed homosexuality as disordered. Homosexuality is not a norm but in terms of statistical occurrence it is a minority condition. However, we have decided as a society that pathologization of homosexuality is wrong. Even if it was the case that homosexuality includes a biological dysfunction, it seems that judging the harm relative to the social standards of a particular time is problematic and even dangerous as it can represent a misuse of psychiatric authority. Sometimes, social standards and values are the ones that need changing, not the individual who can't conform to them.

Another instance of harm in the literature is Rachel Cooper's account of disease. Recall that Cooper is a normativist as she argues that any account of mental disorder is "by its nature anthropocentric and corresponds to no natural class of conditions in the world." (R. V. Cooper, 2002, p. 271). Her view of mental disorder is that it is "a condition that it is a bad thing to have, that is such that we consider the

afflicted person to have been unlucky, and that can potentially be medically treated." (R. V. Cooper, 2002, p. 271). All three criteria are necessary to count some condition as disordered and harm is realized through condition being a "bad thing to have".

In discussing what it means that a condition is a bad thing to have, she dispels some of the common misconceptions about harm in psychiatry. She asserts that "Sometimes it is suggested that something can be a disease if it is a bad thing for society even if it isn't necessarily a bad thing for the potential patient." (R. V. Cooper, 2002, p. 272). These are conditions like personality disorders, pedophilia, and psychopathy. Cooper suggests that this is mistaken since whether a behavior is symptomatic of a disorder does not rest on the behavior itself but on whether the behavior is voluntary. If the behavior is unvoluntary this is a bad thing, both for the individual and the society. Thus, while a behavior might be bad for society, to judge it pathological it must be bad for the person who exhibits the behavior themselves.

Another idea that is suggested is that for something to count as bad it must be such for most of the population, for the majority of typical people experiencing it. However, Cooper argues against this by stating that a particular thing can be a condition even if it was judged by most people as good rather than harmful. She illustrates this with the case of sterility where to people who are sterile but do not want to be, this presents something which is bad for them. However, for people who are sterile and want to be sterile, or even of their own accord engage in procedures such as vasectomy or hysterectomy, sterility is not something bad but wanted. And as such it is not considered to be a disorder. A general principle to extract from this is that "someone can have a disease even if their condition is a good thing for most people" (R. V. Cooper, 2002, p. 272), the most important thing is that it is bad for the person that has the condition.

When it comes to determining what is good (or bad) for some person, Cooper notes that this is a very difficult question, one that plagues the philosophers

throughout numerous philosophical discussions. Determining what is good for people includes a spectrum between what a person actually thinks is good for them, and what an idealized agent or an idealized version of themselves might think is good for them. There are problems with both ends of the extremes:

> To the extent that a method requires idealisation it is obscure how it can deliver answers. I know what I actually value, but how can I know what I'd value if I were more knowledgeable and wiser than I am? To the extent that a method relies on the judgements of actual people it risks giving the wrong answers; after all, actual people make mistakes. (R. V. Cooper, 2002, p. 274)

Bearing this in mind, Cooper goes on to use some common and shared intuitions about what is bad for people without trying to solve this difficult problem. She is adamant, though, that the same condition can be a disorder in one person but not in other – "people have different aims, different abilities and different preferences" (R. V. Cooper, 2002, p. 274), just like sterility is bad for some and good for others. Similarly, some people can be disturbed by hearing voices while others may live healthy and productive lives despite them, or even feel blessed because of their condition.

While Cooper does not elaborate on what really it means that some condition is bad and does not get into what this 'badness' consists in, she provides several important guidelines and principles to adhere to in thinking about harm. First of all, we should concentrate on the harm to the person that has a condition, not the way that they might be harmful for others, society in general etc. We are looking for intrinsic harm of some condition, not the way it may impact various other things indirectly. Additionally, when it comes to conditions we can say that "one man's meat

is another man's poison" meaning that the same condition may affect two people differently where in one it can manifest as harmful while in another it may be neutral or beneficial. While it may seem counterintuitive to some, or even a bit relativistic, I think this is an important insight and an ineliminable component of harm.

Here I have presented the most prominent views on harm in literature. This is not an exhaustive overview of harm in psychiatry as there might be some accounts I have failed to include. This might be considered as an oversight and if that is the case, that is on me. However, I do think it is safe to say that these are the accounts of harm that are most present in the literature, the ones that most authors think about when talking about harm and ones that are addressed in many of the usual cases of harm in psychiatry. While there are some problems with all of these accounts it is hard to formulate harm to fit all of the criteria and do everything we might want from it. Nonetheless, I try to incorporate the virtues of these accounts later in my analysis on harm and hopefully solve some of their shortcomings.

Harm in philosophy of psychiatry is heavily embedded in the debate on mental disorders. There is no separate debate on just harm in philosophy of psychiatry. Rather, harm has conceptual, theoretical and methodological ties to the debate on mental disorders. It is hard to talk about harm in psychiatry outside of the context of the debate on mental disorders, as the debate offers dialectical and conceptual tools to explore the concept and provides us with context in which to analyze it. Thus, in order to explore harm, we need to get familiarized with the basics of the debate, its conceptual apparatuses, discussion dynamic and the methodological intricacies. This I will do in the following chapter.

### 1.4 The Three Pillars of (Good) Psychiatric Practice

The story of depathologizing homosexuality reveals an important insight about psychiatric practice that should be made explicit in our analysis. Observing the course of such a monumental change, we see that there are three important factors of

psychiatry that were brought together and made to communicate with each other. I call these the *three pillars of psychiatric practice* – scientific research, the patient perspective and psychiatric/clinical expertise.

The first pillar, scientific research, encompasses studies in psychiatry and of psychiatric conditions that respect and align to the ethos of the scientific community. It is the "neutral, systematic, planned, and multiple-step process that uses previously discovered facts to advance knowledge that does not exist in the literature." (Erol, 2017, p. 97). Scientific research can be observational or experimental with respect to data collection techniques, it can be descriptive or analytical when it comes to causality relationships, retrospective, prospective and cross-sectional with regards to time and so on (Çaparlar & Dönmez, 2016; Erol, 2017)[5]. It uses the scientific method which "should be neutral, objective, rational, and as a result, should be able to approve or disapprove the hypothesis." (Erol, 2017, p. 97). Scientific research into psychiatric phenomena should, as well as any branch of scientific research, be as valid, objective and impartial as possible while also offering novel data and understandings of psychiatric conditions. Its goal is the advancement of knowledge of psychiatric conditions – the mechanisms behind them, how they operate and function as well as pharmacological trials that aim to investigate treating psychiatric conditions. It includes but is not restricted to research in biology, neurology, psychology, scientific psychiatry and pharmacology.

The second pillar of psychiatric practice is the first person's perspective of the individual that has a particular condition which is comprised under the heading of the patient's perspective. It is giving the voice to the persons receiving health care. Trying to put ourselves in their shoes and examining their life from their own point of view.

The third pillar of psychiatric practice is psychiatric/clinical expertise. This is concentrated in the skills, medical knowledge and experience of clinicians who are in

[5] I will not discuss these at lenght here – see (Çaparlar & Dönmez, 2016; Erol, 2017).

direct contact and provide patient treatment. This may also include clinical psychologists, nurses in psychiatry and social workers which may work in teams with psychiatrists to provide solutions and an all-encompassing care. In this, professional psychiatrists are somewhat authoritative psychiatric experts who have insight into medical theory and practice, psychiatric classifications and symptomatologies of conditions, as well as therapeutic solutions. They are the ones who provide diagnoses and treatment solutions. Their aim is treating patients in psychiatric practice.

All three pillars are and should be to a certain extent independent of each other. Scientific research should be conducted in an impartial and objective way that is not obfuscated or influenced by psychiatric knowledge and personal experiences of particular conditions. Patient's perspective should be approached with similar impartiality, being heard and paid attention to without psychiatrists' suggestiveness and projection onto the patient's experience. Psychiatric expertise should be respected and accounted for as well reserving a certain level of authority to them for making decisions about evaluation and treatment of psychiatric conditions.

Even though they are somewhat independent of each other, these areas heavily influence each other. Psychiatric expertise and care are and should be greatly informed by scientific research in the area. Scientific research of psychiatric conditions should be directed towards solving the problems in the practice and researching conditions that are detrimental to mental health of patients. Psychiatric expertise should be sympathetic and informed about the patient's experiences and their testimonies about conditions.

Good psychiatric practice is supported by all three of these pillars. It depends on all three of these elements working together, interacting, and making compromises. Each should have a stake in coming to psychiatric solutions and each should be respected and represented in psychiatric practice.

Throughout the history of psychiatry, we have unfortunately come to know what happens if any of these elements are not represented in psychiatry. Homosexuality has been considered a mental disorder on the grounds of just one of these elements, psychiatric expertise and authority, while ignoring scientific evidence (or lack thereof) on the issue and while not taking into account the patient's perspective.

Another historical movement that attests to the importance of all three pillars of psychiatric care was the anti-psychiatry movement of the 1960s and 1970s that took place around Europe and the US. The advocates of the anti-psychiatry movement – Thomas Szasz (1960), David Cooper (1967) and Erving Goffman (1961), among others, wanted to give "voice to a growing awareness and protest of the abuses and inefficiencies perpetrated in the name of mental health." (Hamber & Brian, 1996, p. 73). One of the major goals of the anti-psychiatry project was deinstitutionalization. They argued that institutionalization of psychiatric patients was repressive and that human rights of patients were being violated (Hamber & Brian, 1996). The alternative they suggested and fought for was community- and family-based care for the patients. This meant that the patients went from psychiatric institutions to being taken care of by family members and living in the community, with having access to out-patient clinics or some other mental health service (or at least that was the idea). The anti-psychiatry movement succeeded in bringing serious changes and resulted in closing of many asylums and psychiatric institutions for inpatient care across Europe and the US. The patients had finally gained stronghold over their oppressors, the psychiatric and medical authority. However, the grass proved not to be always greener outside the confines of the psychiatric institutions. The issue with "deinstitutionalisation was that communities could not absorb the sudden influx of discharged patients" and that "In most cases, particularly in the West, communities did not have the human resources or infrastructure to cater for this overload."

(Ahmed, 2012; Hamber & Brian, 1996, p. 73). To add insult to injury, the psychiatric patients were often placed back into the environment that caused their mental breakdown in the first place. Many were sent "into under-resourced communities and been abused, stigmatised and marginalised within their very own communities." (Hamber & Brian, 1996, pp. 73–74).

While anti-psychiatric goal was somewhat noble, wanting to ensure protection against misuse of psychiatric authority and giving primacy to the voice of the people, many would argue that it went too far, and it may even have caused more harm than good. The movement wanted to give the power back to the patients but what ended up happening is that many of them suffered worse fates outside of the confounds of psychiatric institutions. Denied the proper care and expertise, they were left at the mercy of their families and communities that might not have had the proper resources or the knowledge to deal with their conditions. These are the dangerous of the 'patient perspective' or the perspective of the users of medical care acting alone without psychiatric expertise and scientific research. It seems that the weight of the treatment of psychiatric conditions cannot be withstood on the patient's perspective alone.

Finally, the history of psychiatry reveals the danger of taking into account scientific research alone, without psychiatric/clinical expertise and the patient's perspective. Especially poignant period of psychiatry to look at is the late 19$^{th}$ century. This is when the biological theories, or at least attempts at explaining mental conditions through biology, really took off (Harrington, 2019). Granted, scientific research on the biological underpinnings of mental conditions were different than that of contemporary research. Their focus was on brain anatomy, they had little understanding of biochemical processes, they were skeptical and 'fatalistic' about treatment of mental disorders (Harrington, 2019). Their fatalism was founded on ideas that some individuals were biologically inept and degenerate and thus

determined to have something wrong with them with little hope of cure or treatment. During such climate the focus completely shifted from patient care to research of their biology. As Anne Harrington asserts "Generations of medical students learned to observe and test patients for evidence of different kinds of breakdown, all while waiting for them to die so the clinical findings could be correlated with autopsy findings." (2019, sec. Anatomical ambitions). This obsession with brain anatomy, and brain dissection spread to asylums like wildfire with the rationale that there is where the diseased brains were concentrated and ripe for the taking. Asylums transformed into research hospitals "outfitted with autopsy theaters, dissecting chambers, and brain banks." (Harrington, 2019, sec. Data from corpses). This was giving off the impression that the patients were more worth, and more interesting, dead than alive. As Anne Harrington quotes from the memoir of Daniel Paul Schreber during his stay in the Leipzig hospital "God does not really understand the living human being and has no need to understand him, because . . . he deals only with corpses." (Harrington, 2019, sec. Data from corpses). This can be considered as a leitmotif of late 19$^{th}$ century psychiatry. Focused on just the biological underpinnings of psychiatric conditions, the researchers of that time have ignored the patient's perspective, seeing them more as brain vessels than persons. They have also forsaken their roles as clinicians focused on patient treatment, replacing it almost entirely with scientism and obtaining scientific data from their corpses. It is evident how inhumane and callous the psychiatric practice may become once we eliminate the patient's perspective and the role of clinicians as care providers.

Thus, all three pillars – psychiatric/clinical expertise, the patient's perspective and scientific research should be equally represented in good psychiatric practice. They should be made to communicate and inform each other with the goal of reaching quality decisions and solutions that would represent a compromise between all of the positions. The history of psychiatry reveals the horrors that might ensue when one of

these three factors is ignored. Learning from these cases is instrumental in knowing why and how to appreciate each of these three elements. This thesis is mostly concerned with one of these elements – the patient's perspective. My aim is to formulate the concept of harm that would most fully and faithfully capture this perspective while at the same time being commensurable with the psychiatric practice and scientific research.

## 1.5 Conclusion

The concept of harm has a rich and longstanding history in psychiatry. It has been introduced in 1970s in the wake of the gay rights movement. In a perfect storm of scientific investigations, activist work and psychiatric interest homosexuality ceased to be considered a mental disorder. Considering this, harm has been introduced as one of the demarcating criteria between pathological and nonpathological conditions. For a condition to be a mental disorder it had to harm the person who had the condition. Since homosexuals are not harmed by homosexuality in itself, homosexuals are not mentally disordered. Harm has been present in psychiatry ever since, being one of the criteria in psychiatric diagnostic manuals such as DSM and ICD. It has also gained prominence in the debate on mental disorder in philosophy of psychiatry. Harm serves the purpose of capturing the patient's perspective which I have argued is one of the three pillars of good psychiatric practice, together with scientific research and psychiatric/clinical expertise.

# 2. The Mental Disorder Debate

## 2.1 Introduction

The central goal of this work is investigating harm in psychiatry. As we have seen in the previous chapter, since its introduction in psychiatry, harm has been considered as one of the criteria of mental disorders. Thus, in exploring harm it is essential to address the debate on mental disorder as this is the conceptual landscape where harm resides. This debate on mental disorders has been central to philosophy of psychiatry for the past several decades and has been attracting widespread attention from psychiatrists, clinicians, philosophers, psychologists, and numerous other stakeholders in psychiatry.

Additionally, there are important consequences that the mental disorder debate has on the social, political as well as interpersonal climate. As the definition of mental disorder draws the line between pathological and nonpathological, the impact this divide has on the individuals, and the collective on either side of it are serious and far-reaching. Thus, I dedicate a portion of this chapter to considering the impact and the meaning that the notion of mental disorder carries in our social practices and individual lives.

In the debate concerning the nature of mental disorders, there are differing views on whether and how harm should enter in their definition and, eventually, how to characterize the relevant notion of harm. This discussion is divided between naturalists and normativists on the opposing sides, with hybrid theories representing a compromise between the two. Naturalists claim that the notion of mental disorder

can be realized exclusively in value-free terms, meaning that the disorder status is a matter of some objective, scientific criterion that is grasped through scientific investigation. Conversely, normativists argue that the notion of mental disorder is essentially and necessarily value-laden, being determined and shaped by social and moral norms we adhere to as a society. Hybrid theorists take elements from both positions fusing them into a single theory.

The notion of harm usually enters from the normative side of the debate, or as an element of a hybrid account since it is mostly viewed as a value term. Harm, however, is not the only normative element but one amongst many contenders on the normative side, together with normalcy, unluckiness, treatability and so on. Harm can also be one among other elements in a theory, being combined with other normative elements in the construction of a purely normative theory, or being an element combined with a naturalist to constitute a hybrid account. I dedicate this chapter to outlining the main positions, arguments, and common objections to each side of the debate. I show how harm places itself in various ways within different theories in the debate.

## 2.2 The uses of notion of mental disorder and their consequences

Mental disorder is a notion that is woven into the institutional, social, political, as well as private fabric of our lives. It is a notion that is used in all walks of life from medical professionals, lawyers, social workers to lay people in the streets, to culture and art - in novels, film, literature and so on. Presumably, all of these operate under some assumptions of what is pathological and have at least some conceptions of the term. Some disciplines, however, that employ the notion of mental disorder show demand for its specificity and precision. There are domains where a well-thought-out division between pathological and nonpathological is critical.

Medicine, or medical practice in general and psychiatry, is one such domain. In deciding what falls under the notion of mental disorder we are signaling which

conditions take priority in medical care. Naturally then, these conditions are also those in which we invest resources in terms of clinical practice, pharmacological and medical research, patient treatment and care. For instance, where we draw the line between pathological and nonpathological has implications for psychiatric epidemiology which aims at determining the rate and prevalence of mental disorders (Horwitz & Scheid, 1999). This in turn impacts mental health policies, justifying funding for mental health research and services, and planning of the efficient distribution of health care (Horwitz & Scheid, 1999).

According to Trautmann, Rehm and Wittchen (2016), when it comes to mental disorders there are direct costs, or "visible costs" associated with diagnosis and treatment such as medication, physician visits, psychotherapy, hospitalizations, etc. Considering that mental disorders are prevalent in the general population and becoming even more so, the costs of mental disorder treatment are beginning to surge. Over 50% of the general population in middle- and high- income countries are going to suffer from a mental disorder at some point in their lives (Trautmann, Rehm & Wittchen, 2016). In deciding what falls under the notion of mental disorder, we direct our resources, and a hefty amount of them, into alleviating these conditions.

However, the weight of carefully delineating pathological from nonpathological does not only lie in the things we invest our resources in but also what is lost by prioritizing some conditions over others. In economics, the term for these are opportunity costs which are defined "as the loss of the benefits the resources could have produced had they been put to their next-best use—the lost opportunity to invest in that alternative." (L. B. Russell, 1992, p. 162). Thus, by deciding that a certain condition should be regarded as a mental disorder, we are steering our resources and attention away from other conditions. Therefore, it is important to choose the right conditions and to choose them wisely to maximize the benefit gained from singling these conditions out.

Especially since when it comes to mental disorders there are also “invisible costs” that can be formulated in terms of income losses due to mortality, disability, care-seeking, lost production due to work absence and so on (Trautmann et al., 2016). To illustrate these invisible costs, consider persons diagnosed with a depressive disorder. They will be considerably challenged in the workforce. For instance, their condition might manifest itself in losses in productivity and efficiency due to psychomotor retardation, tiredness, lack of motivation, all of which are symptomatic of depression. Their condition might also negatively impact their affiliation in the workplace and their cooperation with colleagues, them being a part of a team or working with others, due to the propensity for isolation which is also sometimes a symptom of depression.

Their mental disorder will also impact the lives of their caretakers, partners, parents, children, or other siblings. Their families will have a considerable unpaid workload since the household duties might end up being unequally distributed due to the person’s tiredness, lack of motivation or care for their environment or well-being. This might negatively impact the career or productivity of their family members. It might reduce the opportunities of healthy individuals with whom they are associated either by considerably upping their workload, limiting their freedom to pursue better but maybe impractical work opportunities, discouraging them from going after promotions, and so forth. Thus, it is important to choose the right conditions to be considered under the concept of mental disorder which would minimize these invisible costs by providing care and treatment for the conditions that need it.

Furthermore, the notion of mental disorder stretches into other institutional domains and social practices such as law and the judicial system. One issue that is at the intersection between law and psychiatry are involuntary psychiatric admissions, psychiatric ‘holds’, where patients are kept contrary to their own will to protect their safety and the safety of others. Such admissions are posed as a dilemma between the:

> society's obligation to protect its members by providing care and safety to those debilitated by the ravages of mental illness versus the individual's right to be a self-determining, autonomous agent who is responsible for his/her life choices. (Christensen, 2011, p. 42)

In other words, psychiatric holds pose infringements on civil liberties that prevent some bad consequences such as violence or suicide. It has been indicated that patients admitted in this way usually have a prior history of mental disorder, even up to 95% of the involuntary hold patients that go through the emergency department do (Lachner et al., 2020). In these situations, while the mental disorder status might not be the main and the decisive factor in determining admission and treatment, it is reasonable that it plays a role and might be informative to the staff dealing with the patient. Thus, what counts as mental disorder factors in the courses of action that infringe on civil liberties. It is of grave importance then to pick out the right conditions since the consequences of unrightful and unjustified psychiatric holds can be truly heart-breaking and detrimental to people's lives and well-being.[6]

Another meeting point of psychiatry and law is criminal responsibility. The problem of criminal responsibility lies in determining the criteria, faculties and circumstances under which individuals are culpable for their misbehavior. Some of the questions that arise are:

> Should certain kinds of mental illness absolve people from legal responsibility for their actions? If so, how should the law distinguish between those kinds of disorder that excuse and those that do not? And

[6] A fairly recent example of involuntary detainment in psychiatric institutions is the story of Ana Dragičević, who was kept at Lopača Psychiatric Hospital in Croatia in the early 2000s (Grđan, 2010; McRae & Human Rights Watch, 2010). Although the circumstances of the psychiatric hold seem quite problematic and legally questionable, it is reasonable to assume that the history of unrightfully pathologising homosexuality still permeates both the social and the psychiatric climate in some places.

> how can courts be sure that someone genuinely is mentally ill? (Glover, 1970, p. 102)

The debate on criminal responsibility is longstanding and ridden with psychiatric notions (Craigie, 2015; Glover, 1970; Hart, 2008; Jurjako, 2011; Jurjako & Malatesti, 2018; Malatesti et al., 2020; Turner, 2010).

Besides the institutional and social structures in which the notion of mental disorder is employed, there are further-reaching consequences not immediately apparent, including the effects the diagnosis has on the psyche of the diagnosed, the social upheaval, shunning and stigmatization of the individuals suffering from mental disorder, credibility they are afforded in public, and potentialities for abuse and mistreatment.

One such consequence of diagnosis has been called 'looping effects' (Hacking, 1996). It is a phenomenon in which people who are given a certain label change their characteristics considering this new label, either to adapt to it or to shy away from it. According to Hacking, people classified or labelled a certain way can:

> (…) make tacit or even explicit choices, adapt or adopt ways of living so as to fit or get away from the very classification that may be applied to them. These very choices, adaptations or adoptions have consequences for the very group, for the kind of people that is invoked. The result may be particularly strong interactions. What was known about people of a kind may become false because people of that kind have changed in virtue of what they believe about themselves. (1999, p. 34)

Looping effects are especially interesting because only people are susceptible to them, with regards to psychiatric notions and categories especially, as several examples from psychopathology and the history of psychiatry indicate (Hacking, 1999). One such example is the multiple personality disorder (MPD) which was a

condition diagnosed in the 1980s and 1990s. As the disorder gained a considerable interest from the scientific community as well as from the public the symptoms of the disorder started changing. The explanation was attributed to looping effects as people diagnosed as MPD adopted their symptoms to the perception of the disorder from the public. Additionally, the disorder has been considered a result of child abuse (Tekin, 2012). Consequently, the victims of child abuse would have the multiple personality disorder suggested to them by the psychiatrists after which they would exhibit symptoms of the disorder. Given these controversies, the multiple personality disorder diagnosis has been in large part eliminated and remnants of it have been morphed into what is now called the dissociative disorder.

Moreover, a considerable burden of the mental disorder status is stigmatization. Individuals suffering from mental disorders often face social rejection and exclusion. As Wahl writes:

> Individuals with mental illnesses reported that others avoided them once their psychiatric disorder or mental health treatment was disclosed. Friends, they said, stopped calling, neighbours' visits decreased and social invitations declined, all contributing to an increased sense of isolation and alienation from their communities. (2012, p. 9)

People with mental disorders report of having difficulty making and keeping friends, having close intimate relationships, and keeping ties to their families. It seems that the general population accepts the mentally disordered to an extent in social situations while they askew from more personal relationships with them (Bhugra, 1989). All of this may result in mentally disordered individuals developing self-stigma which is a result of the internalization of the social stigma they are exposed to. As a result, they may believe they are less valued because of their psychiatric disorder (Ritsher et al., 2003).

Stigmatization does not impact just the ones afflicted with the mental disorder, but it stretches out to people with whom they are associated. The stigmatization of persons involved with a mentally disordered person is called "courtesy stigma" (Goffman, 2009) and it affects friends, family and romantic partners. There are numerous stereotypes which facilitate this kind of stigmatization. Some of them are that the public perceives parents as responsible for their child's mental disorder and blames them for it (Corrigan, 2000) and that they consider children or siblings of mentally disordered to be in a way 'contaminated' by them (Phelan, 2005). People shame and blame the individuals associated with those suffering from mental disorders (Corrigan & Miller, 2004).

All of these make apparent the weight of the notion of mental disorder and of the stakes of drawing the line between pathological and nonpathological in the right place. The consequences of doing so badly could be wasted public and private funds, misdirected research and treatment resources, the illegitimate infringement of rights, personal tragedies and ruined life opportunities as well as misdirecting the help away from those deserving and in need of it.

In the following sections I explore how and where the notion of mental disorder is used and what are some formulations of it with a special emphasis on the criterion of harm.

## 2.3 Definitions of mental disorder

The notion of mental disorder is prominently featured in psychiatric diagnostic manuals *The Diagnostic and Statistical Manual of Mental disorders* (DSM) (1980, 1994, 2013) issued by the American Psychiatric Association (APA) and the International *Classification of Diseases* (ICD) (1992) issued by the World Health Organization (WHO). These are considered to be leading authorities when it comes to mental disorders. First and foremost, they are diagnostic tools used to assess and determine the conditions patients might be exhibiting. Notwithstanding, they are also

a reference tool used in various areas of research, treatment, and clinical work. The definition of mental disorder in the DSM – 5, the latest instalment of the DSM is the following:

> A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviant behavior (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above. (American Psychiatric Association, 2013, p. 20)

Let us examine the elements of the definition. We can tease apart two main components of the definition found in the DSM. One component has become widely known as the dysfunction criterion and it rests on the assumption that symptoms which patients exhibit are a manifestation of some underlying dysfunction that has biological, psychological, or developmental underpinnings. The other component in the definition is "significant distress or disability in social, occupational, or other important activities" (American Psychiatric Association, 2013, p. 20), which has become known as the harm criterion.

Let us take note of harm as a criterion that has its place in the definition of mental disorder. It is important to notice, however, that harm in the definition of mental disorder is not a necessary criterion, unlike the dysfunction criterion. As it

says, mental disorders are *usually* associated with distress or disability meaning that it is a common occurrence when it comes to mental disorders but not necessarily so. In other words, there are some disorders which do not involve distress or disability, more generally harm, but are nonetheless disorders.

Distress involves unfavourable internal psychological states in terms of emotional pain, anguish, nervousness and so on, for example in anxiety disorders where a person experiences persistent bouts of nervousness, fear and worry. The other component is disability which involves problems in daily functioning in various presumably important aspects of people's lives. The rationale behind the harm criteria as distress or disability is offered in the DSM – 5 as the following:

> in the absence of clear biological markers or clinically useful measurements of severity for many mental disorders, it has not been possible to completely separate normal and pathological symptom expressions contained in diagnostic criteria. (American Psychiatric Association, 2013, p. 21)

Because of the aforementioned, there are possibilities of persons being symptomatic of a disorder, i.e., presenting with a condition that is not pathological in itself, which would make granting the person a disorder status "inappropriate" (American Psychiatric Association, 2013, p. 21). Therefore, the harm criterion of distress or disability is used to indicate and determine "disorder thresholds" in the form of "the disturbance causes clinically significant distress or impairment in social, occupational, or other important areas functioning." (American Psychiatric Association, 2013, p. 21).

There is significant overlap between the two diagnostic manuals DSM and ICD. In the ICD-10, the notion of disorder is addressed as the following:

> "Disorder" is not an exact term, but it is used here to imply the existence

> of clinically recognizable set of symptoms or behaviours associated in most cases with distress and with interference with personal functions. Social deviance or conflict alone, without personal dysfunction, should not be included in mental disorder as defined here. (World Health Organization, 1992, p. 11)

Here the key elements are shared by both manuals. They contain the element of personal dysfunction, as well as the assertion that the set of pathological symptoms is "in most cases associated with distress and with interference with personal functions" corresponding to the "distress or disability" in the DSM. Again, this is what is referred to as the harm criterion. The definitions of mental disorder are not only the subject of diagnostic manuals but are central to a heated philosophical debate belonging to the philosophy of psychiatry which I outline in the remainder of this chapter.

## 2.4 The Debate on Mental Disorders

Philosophy of psychiatry as a part of analytic philosophy has established itself as a field of inquiry most prominently in the past several decades (R. V. Cooper, 2020; Kingma, 2014). Not exclusive to philosophy, this growing field sparked interested in participants from numerous areas of medicine such as psychiatry, clinical psychology, nursing as well as from other disciplines such as law and social work. The interdisciplinarity and the array of topics discussed among these various scholars are of noticeable academic stature that is growing in relevance.

The central debate in the philosophy of psychiatry is the debate on the definition of mental disorders (Bolton, 2008). The aim of the debate is to find certain characteristics according to which we would be able to delineate pathological conditions from nonpathological. In other words, the idea is to offer criteria which a condition would have to satisfy to be diagnosed as a mental disorder.

Before I start with the outline of the debate, I want to make a couple of terminological points. In this thesis I use the notion of mental disorder interchangeably with notions such as disease, illness, and malady. This is a common move in the discussion (R. V. Cooper, 2002). There are some authors, however, that make distinctions between these terms which I will point out where applicable or relevant. For now, I take these notions to refer to the same thing, or roughly the same thing where the differences are not such that would be pertinent to the discussion itself. Additionally, the notions of disease, disorder, illness, malady are juxtaposed with the notion of health meaning that the conditions referred to as disorder, illness, disease are unhealthy, or they represent some lack of health. Thus, the categories of health and disease are usually taken to be mutually exclusive (Kingma, 2014).

Equally important is to explicate what falls under these notions that I take as synonymous. It is common to consider under the term disorder any kind of pathological conditions (R. V. Cooper, 2002). These include autoimmune diseases, injuries, physical malformations, wounds, with the notion of disorder being used as an umbrella term to cover all these pathological conditions. Therefore, a broken leg, psoriasis, AIDS, gunshot wound, ALS, lymphoma are all considered disorders.

## 2.5 Naturalism

Naturalism is a position according to which the notion of disorder, at least in part, can be analyzed exclusively in value-free, objective terms (Kingma, 2014). It is a position that tries to identify the notion of disorder, in purely scientific terms using the notions such as biological function, brain lesion, and so forth. The naturalist accounts differ in various ways but what they all share is that the notion of disease is a value-free concept (Ananth, 2008; Boorse, 1977, 1997, 2014; Kendell, 1975; Scadding, 1988).

The most prominent representative of naturalism in the debate is Christopher Boorse who has been defending his variant of naturalism for decades (Boorse, 1975, 1976, 1977, 1997, 2011, 2014). On his account the juxtaposition of health and disease

corresponds and can be equated with the dichotomy of normal and pathological (Boorse, 1997).[7] Boorse defines disease as "a type of an internal state of the organism" which:

> (i) Interferes with the performance of some natural function – i.e., some species-typical contribution to survival and reproduction - characteristic of the organisms age; and (ii) is not simply in the nature of the species, i.e., is either atypical of the species or, if typical, mainly due to environmental causes. (Boorse, 1976, pp. 62–63)

Boorse goes on to explain that since diseases are exclusively interferences with natural functions, and since the characteristics of natural function are matters of biological facts, the disease is a value-free concept (Boorse, 1976). This means that there are no value judgments or evaluative claims present in determining whether a condition falls under the domain of disease. Thus, whether something is a disease, namely interference with some natural function, can be determined by scientific means alone (Boorse, 1976).

Notice the centrality of the notion of biological function in his definition, which is characteristic of many naturalist accounts, although not all (Szasz, 1960). However, the notion of biological function used in such accounts is by no means uniform but lends itself to the multitude of different understandings of functions. The debate on functions - on what they are, how they should be conceptualized, is a comprehensive debate in the philosophy of science that overlaps with the debate on mental disorders from the naturalist side (Chin-Yee & Upshur, 2017; Forest & Le Bidan, 2016; Garson,

[7] Boorse even prefers the latter since arguing that health is the absence of disease, as he does, fails to exclude death, meaning that the notions of health and disease are not exactly mutually exclusive (Boorse, 1997).

2012, 2016; Jurjako, 2019). Correspondingly, in the discussion on mental disorders in philosophy of psychiatry views on functions differ just as well (Wakefield, 2014). For the purposes of this work, I will not engage in the discussion on functions lest for roughly sketching the positions to get the sense of what is meant by the notion of function.

Boorse endorses a goal-directedness view of functions according to which organisms aim at certain biologically determined goals, or as Boorse says organisms "adjust their behavior to environmental change in ways appropriate to a constant result, the goal" (Boorse, 1977, pp. 555–556). Now, organisms and their parts have all sorts of goals in which Boorse notices a hierarchy:

> In fact the structure of organisms shows a means-end hierarchy with goaldirectednes at every level. Individual cells are goal-directed to manufacturing certain compounds; by doing so they contribute to higher level goals like muscle contraction; these goals contribute to overt behavior like web-spinning, nest-building, or prey-catching; overt behavior contributes to such goals as individual and species survival and reproduction. (Boorse, 1977, p. 556)

Thus, different functions which organisms have should be viewed in terms of their goals and subsumed under some higher order goals of an organism. Since the organism can have numerous higher and highest level goals, Boorse claims that to some extent "highest-level goals of organisms are indeterminate and must be determined by a biologist's interests." (Boorse, 1977, p. 556), for example, ecologists might consider the highest level goal as achieving homeostasis (Boorse, 1977). For the purposes of medicine, relevant goals are those determined by physiologists and are, according to Boorse, "contributions to individual survival and reproduction" (Boorse, 1977, p. 556).

The biological function is not the only element of Boorse's account of disease. Boorse's account bears the name "biostatistical theory (BST)" emphasizing the two main elements: natural function and statistical normality (Boorse, 1997). He summarizes his account as follows:

> 1. The *reference class* is a natural class of organisms of uniform functional design; specifically, an age group of a sex of a species.
>
> 2. A *normal function* of a part or process within members of the reference class is a statistically typical contribution by it to their individual survival [or] reproduction.
>
> 3. *Health* in a member of the reference class is *normal functional ability*: the readiness of each internal part to perform all its normal functions on typical occasions with at least typical efficiency.
>
> 4. A *disease* [later, *pathological condition*] is a type of internal state which impairs health, i.e., reduces one or more functional abilities below typical efficiency. (Boorse, 1977, p. 562, 2014, p. 684)

However, note that on several places Boorse seems to be open to compromise with regards to the particular view of function he is using claiming the one he vindicates does not have to be the be-all-and-end-all of functions, although he seems adamant about the role of biological function both in his account and in the notion of disease (Boorse, 1997, 2014).

Thus, to determine whether some condition is a pathological one, the individual's reference class (age, sex) should be determined to see whether in that reference class the condition impedes on one's normal functioning. The conditions that negatively influence one's normal functioning are those that hinder the organism's survival and reproduction. When the individual's functioning represents a

statistical abnormality, being below typical efficiency in that reference class, the condition is pathological.

There are several objections directed towards naturalism in general, and Boorse's account in particular. One common objection is that "naturalism does not properly reflect our use of the terms 'health' and 'disease' because naturalism neglects the role values play in determining whether someone is healthy or diseased" (Ereshefsky, 2009, p. 222) (see also Murphy, 2006; Reznek, 1987; Wakefield, 1992).

A well-known example to illustrate this is the case of homosexuality (Ereshefsky, 2009). Homosexuality seems to be a condition that negatively affects one's reproductive prospects. As such, it impedes on the normal functioning of the organism to survive and reproduce. Under Boorse's account, this would mean that homosexuality represents a biological dysfunction, and thus a disorder. The jury is still out on whether homosexuality is a biological dysfunction (R. V. Cooper, 2020), however, consensus that has been gaining ground since the APA's depathologization in 1973 is that homosexuality is not a disorder (and rightfully so).

The circumstances behind the depathologization of homosexuality were not exactly of the nonevaluative, scientific kind reached by new scientific insights into the psychology, biology, or neuroanatomy of homosexuals. The APA's decision was neither initiated nor reached by the biological findings on homosexuality. Rather, it was the determination and relentlessness of gay activist groups and the sympathetic ear of the APA's board towards injustices imparted on gay people that played the key role in the matter (Bayer, 1987).

Furthermore, there are other conditions that seem to involve biological dysfunction but are not considered a disorder. Synesthesia is a condition where "stimulation in one sensory or cognitive stream leads to associated experiences in a second, unstimulated stream" (Hubbard, 2007, p. 193). This manifests in experiences of perceiving letters or numbers as of a particular color, of "hearing colors" meaning

that sounds evoke experiences of colors, and of personification of dates, months, years where such concepts evoke personalities and so on (Hubbard, 2007). Even though synesthesia is considered a neurological condition, and we might infer there being some kind of dysfunction or at least unusual functioning of the perceptive apparatus, we do not consider it a disorder. As Hubbard points out, it is "not listed in either the DSM IV or the ICD classification, as it generally does not interfere with normal daily functioning" (Hubbard, 2007, p. 193). Another similar case includes brain lesions that cause people to crave fine foods (Murphy, 2006).

Another line of objection is that even though naturalist accounts aim at delivering a completely value-free account of health and disease, they fail to do so. These types of objections point to normativity sneaking its way into naturalist accounts. Kingma (2007) presents a convincing case that goes in this direction against Boorse's "biostatistical theory". As I have already mentioned, on Boorse's account to determine whether a condition is a disorder, one must determine whether the condition negatively affects the organism's survival or reproduction relative to the organism's reference class, i.e., age, sex and so on. If a condition statistically deviates from the norm that is determined by a particular reference class, then it is a disorder. For example, to a man, a particular level of testosterone might be expected, even beneficial to their survival or reproduction, thus a healthy state, while the same level of testosterone would be considered to be a disorder in women (Kingma, 2007). The problem, thus, arises because reference classes play a significant role in determining the disorder status of individuals according to Boorse's account.

Furthermore, as Kingma (2007) points out, it is important to determine the *right* reference class. She illustrates that if we were to take heavy drinkers as a reference class, then statistically, what would be a disordered liver functioning in non-drinking population would be a regular, normal liver functioning when measured against the reference class of heavy drinkers (Kingma, 2007). Therefore, she

concludes, it is imperative to be measured against the 'appropriate' reference class to determine someone's disorder status (Kingma, 2007).

Now, here is the crux of her argument. Although Boorse claims that the matter of health and diseases rests entirely on factual matters, empirical facts, Kingma brings that into question by wondering what determines which reference classes are appropriate (Kingma, 2007). The reference classes could have been arranged and chosen in any number of ways which would give different results with respect to the statistical normality, and consequently the disorder status (Kingma, 2007). Boorse does not seem to offer a satisfying empirical justification for choosing one reference class over another on his account. Therefore, she concludes:

> because the choice of reference classes determines the distinction between health and disease on the BST, and Boorse gives no empirical fact that justifies the choice of these reference classes over others, there is no empirical fact that determines the distinction between health and disease on his account. (Kingma, 2007, p. 131)

Thus, given that no empirical fact determines the disorder status, naturalism fails at delivering what it sets out to do, to offer a value-free, completely factual, account of mental disorder.

Kingma's reference class objection to naturalism, as well as the aforementioned objections, pose significant problems for naturalism as well as for settling the disorder status of the conditions that are in the purview of psychiatric interest. Let us now consider the normativist side of the debate.

### 2.6 Normativism

Normativism with regards to the definition of mental disorder is a position according to which the notion of disorder is necessarily, and in its essence, value-laden (Agich, 1994; Bolton, 2008; Clouser et al., 1997; R. V. Cooper, 2002; Engelhardt, 1976; Fulford,

1989; Nordenfelt, 1995, 2007; Reznek, 1987). This means that the concept of disorder can never be fully realized or defined in value-free terms but is necessarily evaluative. Granted, normativism does not necessarily entail that there is no place whatsoever for biological or some kind of naturalistic factors. Some factors can be described and explained in naturalistic terms. But these are not what make mental disorders.

The appeal of normativism lies in that our views of what we generally deem normal, usual, and commonplace depends largely on how societies function. Take dyslexia, for example. Since as a society we put high value on the skills of reading and writing, it being ubiquitous in all walks of life, we deem inability or seriously subpar performance in these skills to be problematic. Because of this, we tend to seek ways of alleviating and treating such problems to improve the quality and opportunities the people with such conditions have. If we imagine a society that does not put emphasis on reading and writing, then dyslexia ceases to be salient, and some other condition could take its place.

One such normativist account is Rachel Cooper's. According to her, a disorder is "a condition that it is a bad thing to have, that is such that we consider the afflicted person to have been unlucky, and that can potentially be medically treated." (R. V. Cooper, 2002, p. 271). That diseases are bad things to have is evocative of the harm criterion in mental disorders which I have pointed out earlier. Specifically, it means that "a condition can only be a disease if it is a bad thing for the potential patient" (Cooper, 2002, p. 272). In contrast to the previously mentioned naturalist accounts, Cooper goes on to argue that biologically different characteristics are not enough to count as disordered but that the difference would have to be harmful to the individual (R. V. Cooper, 2002). In line with this, she mentions ginger people and highly intelligent people which are different, we could presuppose even biologically different, but are not disordered (R. V. Cooper, 2002).

Cooper adds that besides a condition being a bad thing to have, a person

should be considered unlucky to be afflicted by it. Moreover, the condition must be potentially treatable (Cooper, 2002). The notion of unluckiness captures the expectedness or unexpectedness of certain condition with regards to the person's age, sex and the general state of the average individual of their demographic. To illustrate this Cooper mentions how male baldness is not considered a disorder while female baldness is (Cooper, 2002). Another example is loss of body hair in various parts of body which seems to happen expectedly with age, however, when it happens rapidly and in young people, it is a cause of concern being either a serious symptom or a disorder known as alopecia. In this regard, the idea of unluckiness is like Boorse's notion of statistical normality that I have mentioned earlier, or statistical infrequency found in Kendell (Kendell, 1975).

Lastly, the element of the potentiality of treatability immerses the account into the abilities and pragmatics of the medical practice. According to this criterion, for a condition to be considered a disorder it must be potentially treated by medical sciences meaning that while there might not be an existent treatment, we could reasonably hope or anticipate the discovery of it (Cooper, 2002). Bad and unlucky things are not enough to be considered medical conditions, for example poverty or homelessness, but bad and unlucky conditions can find their way into medicine once they are considered appropriate to address by medical practice and once, they become treatable or manageable by medicine (Cooper, 2002). This criterion does the work of biological dysfunction in naturalist accounts meaning it grounds the criteria in the practice of psychiatry, but Cooper (2002) wants to steer away from biological basis for the notion of disorder for the following reasons.

Cooper (2002) considers relying on the assumption that mental disorders have a biological basis both too strong and too weak. She considers it to be too strong because there might be some mental disorders for which we would not be able to find a biological or a neurological substratum (R. V. Cooper, 2002). She also considers

the criterion of biological basis to be too weak which she illustrates with the following:

> Having a bad haircut and being unable to fit into last year's clothes are bad things, sufferers may be unlucky, and both have a biological basis, but they are not diseases. They are not diseases because we do not rely on medical help to fix these problems. (R. V. Cooper, 2002, p. 277)

Thus, instead of the biological basis, she opts for the potential of treatability as one of the criteria she endorses in her account of mental disorder. Cooper's account which we have analyzed so far is normative in that the criteria of badness, unluckiness, and the potential of treatability, while may dip its toes into some naturalist elements remains in large part value laden.

Just like naturalism, normativism is subjected to numerous objections. One of the problems, pointed out by Dominic Murphy (2006) strikes at the very heart of normativism. It argues that normativism fails to capture the way we think about mental disorders - it "cannot explain our judgements"(Murphy, 2006). As I have previously mentioned, the essence of normativism is that mental disorders are determined and shaped by social and moral norms alone, rather than resting on a failure of some biological or psychological functions. However, if mental disorders are states we simply disvalue, how do we then differentiate these from other states we disvalue? If violating social norms or deviating from them counts as a mental disorder, why don't we consider all deviations from the norm as pathological? (Murphy, 2006). As Murphy (2006) emphasizes, we disvalue things like racism, cowardice, hypocrisy but we seem to clearly make a distinction between these states and pathological conditions. There are also behaviors or personalities that deviate from social norms as probably any of us knows someone whom they could characterize as eccentric. Nonetheless, we do not consider such characterization inherently tied to the pathological. This might be indicative that we ground our judgements of what we

consider as pathological in something other than just our value-judgements. It seems that resting the notion of mental disorder entirely on things we disvalue, or which deviate from social norms is insufficient and does not tell the whole story of how and what we consider pathological.

Another problem with normativist accounts lies in the fact that society can be wrong in its evaluation of what is normal, appropriate and/or pathological. Unfortunately, we have the notorious history of psychiatry to attest to the wrongful and unjust treatment of socially condemned and marginalized groups. Take the example of drapetomania. At one point, Samuel Cartwright, a physician in the American South in the nineteenth century postulated a 'disorder' that afflicted slaves who exhibit the behavior of running away from their owners (Cartwright, 1851). Another such example is hysteria, a 'disorder' that afflicted primarily women, caused by the female reproductive organs, which manifested in sadness, anger, insubordination and so on (Ehrenreich & English, 2013). There are many other dubious conditions whose disorder status was from today's perspective obviously unrightful and unscientific for example, masturbation (Engelhardt, 1974), lack of female orgasm (Kaplan, 1983) and so on. These cases warn about the dangers of the unquestioned psychiatric authority inadvertently influenced by the societal values, as well as shake the ground on which normativism stands as a theoretical approach to the notion of mental disorder.

## 2.7 Hybrid Accounts

Hybrid accounts combine elements from both groups of theories into a single account. They present a compromise and aspire to settle the problems that rid each of the polarized positions. One prominent and most promising hybrid account of the disorder is the Wakefield's harmful dysfunction analysis (HDA) (Spitzer & Wakefield, 1999; Wakefield, 1992, 1997c, 2006, 2007, 2014; Wakefield & Conrad, 2019, 2020). On this account mental disorders contain two criteria, harm, and biological

dysfunction, both of which must be satisfied for a condition to count as a mental disorder (Wakefield, 1992). According to Wakefield:

> A condition is a disorder if and only if (a) the condition causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion), and (b) the condition results from the inability of some internal mechanism to perform its natural function, wherein a natural function is an effect that is part of the evolutionary explanation of the existence and structure of the mechanism (the explanatory criterion). (Wakefield, 1992, p. 384)

Thus, only biological dysfunctions that are harmful are disorders. Things that are harmful but are not biological dysfunctions are not disorders and things that are biological dysfunctions but not harmful are also not mental disorders. The account follows the intuitive cases of what we regard as mentally disordered while avoiding some of the problems on naturalist and the normative side. It purports to ground normativism as to avoid the 'anything goes' objections of social relativism and it provides naturalism with additional, nuanced criteria to build onto the 'cold objectivism' of value-free biological dysfunctions. Homosexuality, regardless of whether it is a biological dysfunction, is not harmful to the individual and as such, not a mental disorder. Cowardice, hypocrisy or being unattractive, even though disvalued by a person's culture and thus harmful, presumably does not have an underlying biological dysfunction, therefore, these things are not a mental disorder.

Let us look closely at each element of the definition starting with the naturalist element of biological dysfunction. As I have previously indicated, (b) the biological dysfunction is the criterion belonging to the naturalist accounts, similar to Boorse's. Even though both Wakefield and Boorse take biological dysfunction to be a necessary component, their views on what biological notion entails significantly differ. In

contrast to Boorse's idea of dysfunction in which anything that impedes on organisms' survival and reproduction counts as one, Wakefield's starting point is evolutionary theory. According to Wakefield, dysfunction is a "failure of a mechanism to perform its natural function" (Wakefield, 1992, p. 382). Now, a natural function that Wakefield takes as relevant is the one that the organ is evolutionary 'designed' to perform. Organs have many effects, but it is important to tease apart the ones that perform a natural function for just any effects the organ has. For example, the sound of the heart performs a useful function in medical diagnosis, however, we do not take this as relevant from the evolutionary perspective, rather the function of the heart to pump blood is the role that the heart naturally plays in an organism, the cessation of which causes serious dysfunction (Wakefield, 1992).

This analogy extends to mental faculties as well, for example, memory plays an important role in learning which seems to have a role from the evolutionary perspective in organism's survival and reproduction. We seem to categorize learning impairments and memory impairments as disorders in part because they fail to do what they were presumably selected for to do and which seem to be of considerable importance in the evolutionary history of the species.

However, Wakefield's notion of biological dysfunction based on evolution faces criticism. One such problem is that from the evolutionary perspective, it does not seem to be necessary that dysfunction must be involved in behaviors or mental states which would be of psychiatric interest. Important factor in any kind of evolutionary explanation is the relation between the organism and the environment. As it goes through the motion of evolutionary processes such as migration, random drift and mutation (Jurjako, 2019), circumstances which would make one trait adaptive and desirable from a perspective of survival and reproduction might turn out to be maladaptive and detrimental to the organism's survival.

Woolfolk and Murphy (2000) offer a useful analogy to illustrate this point.

Putting a smoke detector near a stove is a sure way to make the smoke detector go off on many occasions of using the stove. Now, if you were to assume that the problem is in the device itself and take it apart to find the bug or malfunction in the mechanism of the device, you would be wrong and not have much luck in finding the problem in the device. Rather, as Woolfolk and Murphy (2000) say, the problem is "environmental" meaning it is in the dynamic between the device and its surrounding. The device might be working precisely as it was intended to do, but it may be placed in the environment, where working according to its design is not very functional, informative, or helpful.

Analogously, some mental mechanisms or functions might be working just as they were selected for, reflecting their proper function and not a dysfunction. However, the environment might be such that the trait or a function that was once selected for might have become evolutionary disadvantageous. As Woolfolk and Murphy write:

> Mental disorders characterized by fear and avoidance may be cases in which internal mental mechanisms are functioning as designed. Identification and avoidance of dangerous situations is adaptive, and the "better safe than sorry" strategy is, relative to suffering false negatives, adaptively superior across the spectrum of species. (2000, p. 244)

So, while it may seem that some conditions are dysfunctions as failures of some mechanism to perform its natural function, they might not be a consequence of dysfunction at all but the individual might be exhibiting a trait that was evolutionary selected, adaptive for some environments but just not very adaptive in the current environment (Jurjako, 2019; Murphy & Woolfolk, 2000). For instance, psychopathic traits, which we compellingly regard as pathological in the way the society is currently organized, might actually be adaptive in harsher, more dangerous environments in

which there is shorter life-expectancy (Jurjako, 2019). Thus, biological dysfunction might not be involved in states and behaviours we consider of psychiatric interest and which we deem pathological, which might make its status with regards to the notion of disorder questionable.

In addition, Bolton (2008) emphasizes another important problem when it comes to the Wakefield's biological dysfunction element (see also (Graham, 2013). The problem is of the practical nature where even if we take Wakefield's evolutionary view of dysfunction as a promising scientifically based endeavor to ground disorders, applying it to the clinical practice leaves something to be desired. According to Bolton, it would then be:

> impossible to make a definite diagnosis of mental disorder in the clinic, because diagnosis would be conditional on a hypothesis that the presenting problem is or involves failure of an evolutionary designed mechanism to function as designed, a hypothesis that would typically be, for most psychiatric conditions, uncertain, speculative, provisional, for some quite likely false – and in probably all cases controversial. (2008, p. 20)

Thus, instead of strengthening the scientific robustness and validity of psychiatric diagnosis by grounding it in a biological element, which the naturalist elements, or naturalist accounts purport to do, it would bring into the definition of mental disorder even more questionable and shaky elements, which the naturalist element purports to avoid.

Let us leave the notion of biological dysfunction aside and focus now on the harm criterion. According to Wakefield, for a condition to be a mental disorder, the condition "causes some harm or deprivation of benefit to the person as judged by the standards of the person's culture (the value criterion)" (Wakefield, 1992, p. 384). Even

though biological dysfunction is a necessary criterion, it is not sufficient as "only dysfunctions which are socially disvalued are disorders" (Wakefield, 1992, p. 384). Wakefield (1992) recounts instances of a dysfunction which are not harmful to the individual, for example, albinism, situs inversus (reversal of the heart's position), fused toes and so on. He concludes that even though they are a dysfunction of some mechanism, they are not disorders, since they are not harmful to the individual as deemed from the society's perspective.

The harm component is not without its objections. An underlying idea behind the discussion of disorder, especially on the naturalist side where biological dysfunctions are a focal point or containing a naturalist element like Wakefield's account does, is that the concept of disorder should be applicable to all organisms, not only humans (Boorse, 2011). For example, a white moth in a sooty environment is considered disordered (Boorse, 2011). However, Boorse (2011) points out that there is 'little social concern' and no cultural standards for moths meaning they would not satisfy the harm requirement and not be considered disordered on Wakefield's account.

Feit (2017) finds fault with the fact that Wakefield defines harm in relation to the social standards of a person's culture. He points at the problem of cultural relativism[8] in such a view, which brings problematic and counterintuitive results (Feit, 2017). For instance, society is not uniform in its values meaning that there might be opposition and ambivalence with regards to values among different groups (Feit, 2017). Additionally, some subgroups or subcultures might subscribe to values and cultural standards that might significantly differ from the majority of the society. Feit (2017) provides the example of pro-ana communities where anorexic standards of

[8] „Cultural relativism is, roughly, the view that social values are what make any given action right or wrong, so that cultural differences in values imply differences in the moral status of the very same type of action across cultures." (Feit, 2017, sec. Against the social-values account of harm)

appearance are celebrated and the norm while in the rest of the society they are the exception. Another such example can be deaf communities where being deaf is the standard, and according to some they consider themselves to be a “cultural minority group” rather than persons with disabilities (Sparrow, 2005).

In addition, judging whether something is a disorder according to the standards of a certain culture means that the same condition intrinsically can result in different disorder statuses in different cultures as the same condition would be regarded as harmful in one, but not in the other culture (Feit, 2017). For example, schizophrenia in our culture is considered a disorder, but we can imagine a culture where a person with schizophrenia might be regarded as a spiritual leader that converses with spirits.

Thus, Wakefield’s account has been criticized on both elements consistent with the tug-of-war atmosphere between naturalist and normativists in this debate.

## 2.8 Conclusion

The notion of mental disorder has numerous applications and is present in all walks of life, both public and private. The most considerable impact it has is in medicine where the scope of the notion of disorder has consequences on resources distribution and allocation, on devising treatment and pharmacological research as well as on patient care. Outside of medicine, the notion of disorder is also a factor in judicial system where it plays a role in infringements of rights such as the involuntary psychiatric holds and in the questions of criminal responsibility in the court of law.

Furthermore, it affects the ones diagnosed with the disorder as well as those in their lives as mental disorders carry with them hardships and suffering connected to the condition itself as well as the stigmatization attached to mental disorders.

All of these illustrate the importance of defining the notion of mental disorder and defining it right. It emphasizes drawing the line between pathological and nonpathological in exactly the right place to maximize the resources allocation in

medicine, to treat the conditions in most need of treatment and to alleviate the suffering of the individuals and their families.

In this thesis I shed light on the notion of harm as one of the aspiring contenders for the definition of mental disorder. To explore the notion of harm with regards to mental disorders it is of essence to engage the discussion on mental disorders in philosophy of psychiatry. The discussion on the definition of mental disorder is ongoing with both naturalist and normativists fighting to gain the upper hand, and hybrid theorists offering various nuances between the two camps. Naturalists argue for a completely value-free notion of mental disorder often in terms of biological dysfunction. Normativists consider mental disorders to be an essentially value-laden category. Harm enters the debate from the normative, or from the hybrid side of the debate. As I have outlined in this chapter, both positions strike an intuitive chord with what we think mental disorders are and should be. However, both side of the debate face numerous objections as well.

As we will see in the next chapter, some authors started to question the fertility of this debate on mental disorders in philosophy of psychiatry. Some, grown disheartened by the stalemate that has been persisting for decades, started to question whether the participants in the debate are going about it the right way. Thus, a metadebate started to emerge that has in its focus the methodology of the original debate. I dedicate the next chapter to engaging with this debate and bringing into it some metaphilosophical insights.

# 3. The limits of the 'traditional' debate on mental disorders

## 3.1 Introduction

In the previous chapter, I have outlined the debate on mental disorders that has persisted for the last forty years. This debate is considered as traditional and central to the philosophy of psychiatry. It is a starting point in our investigation of harm as well.

Recently there has been a significant pushback against the traditional debate. The long-windedness paired with the lack of conclusive results lead some authors to proclaim debate to be in a stalemate. Several authors have brought into question the content, as well as the methodology of the traditional debate on mental disorders. One of the most contested elements of the traditional debate has been the method of conceptual analysis. Yet, there are also other elements that are brought into question like the nature of the concept at hand, whose intuitions and why should be considered, the different scopes, intention and extensions of the concept that main contenders are investigating and so on.

These objections are not exclusive to the debate on mental disorders. Rather, they are reflective of broader revisionary attitudes many contemporary philosophical authors take with respect to contemporary philosophical discussions. There is an atmosphere in the contemporary philosophy that challenges the traditional precepts, methodology and topics, a practice that is becoming increasingly important and spilling over into much of the contemporary debates. As a result, we witness a rise in

popularity and importance a field of philosophical inquiry that examines philosophy itself called metaphilosophy.

Being aware and well-versed in metaphilosophical literature and ideas allows one not to just do philosophy but think, explore and be creative in what, how and why one is doing philosophy. Since many of the metaphilosophical insights and ideas are present in numerous contemporary debates, it is both wise and responsible to engage them. Thus, in this chapter I point out important metaphilosophical points, mostly focusing on the debate on mental disorders. We will see the move from the 'traditional' debate using conceptual analysis to novel methods and modes of inquiry. I take these to be just as relevant in our investigation of the concept of harm. They will serve as methodological building blocks to the notion of harm which I will be developing in the following chapters.

This chapter is an effort to systemize and evaluate the objections to the traditional debate. I argue that the grievances directed at the traditional debate are reasons enough to consider alternatives in thinking about both the notion of mental disorder as well as the harm criterion.

### 3.2 The critique of the traditional debate on mental disorders

In recent years, the literature on the objections directed at the debate on mental disorders has been steady accumulating. Even more so, the lack of dynamic in the debate has been characterized by some as a "dull thud of conflicting intuitions" (Bigelow & Pargetter, 1987, p. 196; Schwartz, 2014, p. 576). The issues are manifold ranging from the objections directed to the methodology of the project, namely conceptual analysis, to the scope, purpose and the utility of the concept of mental disorder.

Harold Kincaid shortly summarizes some of the central elements that are brought into question:

> That project has proceeded in the traditional style of definitions and

> counterexamples, based on what we would call a disease or would not. The problems with this project are many. Who is the "we"? Any reasonable person? Any physician? Any pathologist? Why think there is any one concept that is capturable by necessary and sufficient conditions, when most of our concepts do not come in this form in the first place? What would we learn if we could complete the project? (Kincaid, 2008, p. 370)

Before I go into these elements, there are tentative positions one can assume in this reformatory climate. Specifically, there are several ways in which the authors who contest the traditional debate explain the lack of results of the debate and the presumable stalemate of it. One strand proposes that there is still hope, meaning that the definition of health and disease really exists but is hard to find. Halvor Nordby (2016) dubs these the realists about the existence of the definition of mental disorder. The realism consists in the idea that they believe that the definition of mental disorder is 'out there' irrespective of our failed efforts at grasping it. According to this view, it exists and presumably can be grasped. However, our efforts so far have not been successful in capturing it. Worrall and Worrall (2001, p. 49) capture this line of thought as such: "the explicatory research programme may presently be degenerating, but it might be destined to "stage a comeback" and perhaps even find the *correct* definition of disease.". Considering the longevity and the lack of conclusive results, we might as well call them optimists.

The opposite, more skeptical and pessimistic strand of thought argues that the lack of success in finding the definition of mental disorder is reflective of the fact that there is none: the definition of mental disorder cannot be found, at least not in the way that the search has been going so far. Worrall and Worrall are partisan to this view, ending their analysis of the debate with the following:

> since there is no reason to think that any such adequate general

> characterisation of disease can be developed, this is not a sensible way in which to approach the issue of the status of mental conditions. (2001, p. 55)

Hofmann (2001, p. 230) shares their sentiment at least in part by claiming that "Disease is basically an issue that is so complex that it appears extremely difficult to encompass it by a single monistic theory.", concluding that it "does not lend itself to a simple definition."(Hofmann, 2001, p. 230).

Nordby (2016) is a proponent of the pessimist view as well. He goes a step further and argues that while Hofmann and Worrall and Worrall offer inductive arguments for the unlikeliness of the definition of mental disorder[9], Nordby (2016) gives an account on why it is in principle impossible to give an account of mental disorder. This he bases on the philosophy of language, following a famous paper by Quine "Two Dogmas of Empiricism" (1976). In it, Quine advocates against the divide between synthetic and analytic propositions and argues against the attempt to reduce meaningful statements in a language to logical constructs. I will not go into detail regarding Nordby's arguments against the possibility of coming to a definition of mental disorder. His argumentation, even though pointed at the notion of disorder, actually extends much further showing that "our English concepts do not have definitions."(Nordby, 2016, p. 177). For the purpose of this work it is suffice to say that he thinks it is "impossible to define concepts in the way one has attempted to analyze disease" because such endeavors "must either be question-begging or viciously circular." (Nordby, 2016, p. 174) (for a complete argumentation see Nordby (2016)). This is a general argument that Quine puts forth in his paper and Nordby applies it to the discussion on mental disorders. Even though Nordby offers an

[9] Inductive arguments are those based on the fact that 'so far' we have not been successful, that data so far shows us that it has not been tenable. (for the discussion see Hawthorne, 2021)

interesting and an elegant analysis in application of 'Two Dogmas of Empiricism' I feel that the argumentation is much too general and extends much further from parameters of our meta-philosophical considerations about the debate on mental disorders to be engaged with at length.

Thus, some authors claim that there is still hope in coming up with a comprehensive notion of mental disorder, those being realists or optimists with regards to the results of the debate. On the opposite side, there are authors who are skeptical or rather pessimistic of the efforts of the debate. While the views and evaluations of latter authors range from careful agnosticism to downright pessimism about the results of the debate, I subsume them all under the general umbrella-term of skepticism.

It is hard to generally align oneself with either the realists or the skeptics without knowing which elements of the debate are in question, which efforts, and what we actually want from the notion of mental disorder. In this chapter I show that there are numerous problems with the traditional debate that render these positions, of either skepticism or realism, not fruitful enough. Whether or not the notion of 'mental disorder', and the notion of harm, is somewhere out there to be discovered is asking the wrong question. Especially, concerning all of the things that are unclear and obscure in the debate. Under what conditions and under which parameters does it make sense to talk about the notion of mental disorder and the notion of harm? – that is the question that it makes sense to ask. In the following, I analyze and assess some of the elements that are characteristic of the debate. I outline the problematic instances and points of disagreements while emphasizing the lessons that can be drawn from these problems.

Let us, thus, address the issues Kincaid merely touches upon one at a time starting with one of the most contested elements of the debate – the method of conceptual analysis.

### *3.2.1 Conceptual analysis*

So far, the discussion on mental disorders between naturalists, normativists and hybrid theorists relied on the method of conceptual analysis (Lemoine, 2013; Walker & Rogers, 2018). The method of conceptual analysis is the traditional method of analytic philosophy and the method can be traced back to the birthplace of Western philosophy, the Ancient Greece. There you could see Socrates, a figure many consider to be the father of Western philosophy, walking around Athens, asking citizens questions about the concepts such as justice, truth, knowledge, virtue, piety and so on (Plato, 1997).

These conversations would start by Socrates asking a question of the form 'What is X?', for example, X standing for courage, piety, virtue, justice, knowledge and so on. The conversation would continue by his interlocutor providing a definition, a way of looking at a certain concept Socrates inquired them about. Then, Socrates would question the interlocutor further, suggestively offering cases where the interlocutor's initial definition falls short. The result would be the interlocutor's further refining his definition, and again, Socrates offering cases to test it. This would go on until the conversation would come to a halt, known as the *aporia*. This became known as the Socratic dialogue, or the Socratic method of inquiry, and is a method spread both in and widely outside of the philosophical confounds.

In the contemporary analytic tradition, at least two thousand years after Socrates, conceptual analysis is described as the following, it purports to:

> (…) think up a possible general characterisation of the cases falling under some concept C and then to test it by trying to find or imagine a particular situation which fits the suggested characterisation and yet would *not* be a situation to which C could be truthfully applied. (Overgaard et al., 2013, p. 85)

In observing this definition, we can see striking similarities between it and the Socratic method. The interlocutor would offer a definition, i.e., a general characterization of cases falling under some concept C, and Socrates would test this by offering counterexamples. The goal of this process would be to specify the conditions in which the concept holds or encapsulate the set of cases that the concept is applied to. In other words, we want to see what are the elements that always remain present in the application of the concept. This is often referred to as searching for necessary and sufficient conditions in the philosophical jargon (Strawson, 1992).

A famous part of Plato's dialogue Meno beautifully illustrates this. Socrates asks Meno to provide a definition of *virtue* (Plato, 1997, p. 872). Meno starts by describing and enumerating virtues of a man, virtues of a woman, virtues of a child claiming that there are many virtues concluding with "There is virtue for every action and every age, for every task of ours and every one of us" (Plato, 1997, p. 872). Socrates humorously retorts to this the following, and this is how the dialogue continues:

> SOCRATES: I seem to be in great luck, Meno; while I am looking for one virtue, I have found you to have a whole swarm of them. But, Meno, to follow up the image of swarms, if I were asking you what is the nature of bees, and you said that they are many and of all kinds, what would you answer if I asked you: "Do you mean that they are many and varied and different from one another in so far as they are bees? Or are they no different in that regard, but in some other respect, in their beauty, for example, or their size or in some other such way?" Tell me, what would you answer if thus questioned?
>
> MENO: I would say that they do not differ from one another in being bees.
>
> SOCRATES: If I went on to say: "Tell me, what is this very thing, Meno, in which they are all the same and do not differ from one another?" Would

you be able to tell me?

MENO: I would.

SOCRATES: The same is true in the case of the virtues. Even if they are many and various, all of them have one and the same form which makes them virtues, and it is right to look to this when one is asked to make clear what virtue is. (Plato, 1997, pp. 872–873)

As we can see, upon asking what virtue *consists in* Meno responded by giving him various examples of virtues. Socrates corrects him by saying that he got a whole swarm of virtues when in fact Socrates inquired about what virtue is in itself. Drawing on the example of bees Socrates explain how bees tend to differ in their characteristics – they can be bigger, smaller, more or less beautiful, but what they do not differ in is the fact that they are all bees – that is something that is similar and shared by all the bees alike. Socrates supposes that there is something which all bees share, the thing that makes them bees, *bee-ness* if you will. Applying the analogy to virtues, Socrates is looking for the things that are shared by all virtues, by all of these particular instances of virtue that Meno enumerates. He is looking for the characteristics that make virtues the thing that they are.

To do this, he is turning to conditions that have to be satisfied for the application of the concept. In the philosophical tradition, we call these the necessary and sufficient conditions. Necessary conditions are those that *have to* be present (i.e. true) for the application of concept: "If p is a necessary condition of q, then q cannot be true unless p is true"(Blackburn, 2008, p. 71). Sufficient conditions of a given concept, if met, guarantee the fulfilment of the application of the concept: "If p is a sufficient condition of q, then given that p is true, q is true as well." (Blackburn, 2008, p. 71). For example, to open the window in my office I have to pull the lever down and towards myself. That is necessary for me to open the window (instead of breaking

it) because otherwise you cannot get it open. It is also the sufficient condition of opening the window as that is all I have to do to open it. If I do this, and just this, I will open the window. This is a situation in which the necessary and sufficient conditions are satisfied by the same thing. There are also other combinations where the condition is necessary but not sufficient, where it is sufficient but not necessary and all the combinations between them. For example, a necessary condition of being a professional guitarist is being able to play the guitar. However, being able to play guitar is not a sufficient condition as many people who play the guitar are not professional musicians. Thus, in looking for the definition of bees, as well as in looking for the definition of virtue, one seeks to find necessary and sufficient conditions for the application of the concept.

Characteristic traditionally ascribed to conceptual analysis is that it is an a priori endeavor. *A priori* usually refers to the justification of knowledge and is contrasted with the *a posteriori*. When we know, or come to know some proposition a priori it means we come to that knowledge just in the virtue of understanding of that proposition, just by thinking about (B. Russell, 2020). Thus, by thinking about the proposition "Triangle has three angles" we come to know it is true simply by understanding what 'triangle' means. The same goes for "all bachelors are unmarried men", by knowing what 'bachelor' means we come to know whether the statement is true or false. We do not need experience from the outside world to know whether these propositions are true but we can come to their understanding from an "armchair"[10]. Conversely, a posteriori knowledge or justification of knowledge

requires more than merely understanding a proposition. Observations

[10] A priori kind of thinking and judgements have come to be known as "armchair thinking" as we do not need to leave the armchair, i.e. go out and empirically explore the world to come to some understanding or knowledge about such propositions.

> based on our senses, or introspection about our current mental state, are needed for us to be empirically, or *a posteriori*, justified in believing that some proposition is true. (B. Russell, 2020, sec. 1. Examples that illustrate the difference between a priori and a posteriori (empirical) justification)

This means that the *a posteriori* includes empirical experience of the world and propositions have to be checked against the state-of-affairs in the world. For example, propositions like "Triangles are the most common shape in the molecular structure of the world", "Bachelors tend to live longer than married men". The dichotomy between a priori and a posteriori is common in philosophy and permeates many discussions. Generally, philosophy has traditionally been seen as an a priori discipline. Many see philosophy as "essentially the a priori analysis of concepts, which can and should be done without leaving the proverbial armchair." (Margolis & Laurence, 2021, sec. 5. Concepts and Conceptual Analysis).

Another characteristic of conceptual analysis is that it is a descriptive project. This means that it explores the concept of mental disorder as it *is.* It seeks to encapsulate the application of the concept in various context to find regularities in its application. This leads to development of a definition in terms of necessary and sufficient conditions. The results of such analysis can tell us how the concept is used and what linguistic community thinks about what the concept represents.

Returning to the notion of mental disorder, the standard discussion was mostly carried out as such, being a descriptive project dedicated to closing in on the necessary and sufficient conditions. This can be observed in leading authors in the debate which engage in providing and refuting counterexamples (Boorse, 1997, 2014; Wakefield & Conrad, 2020). Lemoine (2013) notes that these attacks are basically of these three forms:

> They consist in presenting (1) cases falling within the commonly accepted extension of the term but which do not satisfy the opponent's definition, (2) cases that do satisfy the opponent's definition but which fall outside the commonly accepted extension, and (3) cases that fall clearly inside or outside the extension but which the opponent's definition fails to classify at all. (2013, p. 310)

When it comes to mental disorders, this means: 1) finding cases we normally think of as mental disorders but are not included in a particular definition, 2) examples of mental disorder that fit some definition which we don't normally consider as disorders and 3) mental conditions that may or may not be cases of mental disorder but which the definition fails to address.

Lemoine (2013, p. 311) contends that the major and prolific authors of the debate, Nordenfelt, Boorse and Wakefield, follow the rules of the conceptual analysis in that they:

> (1) propose a definition of health, disease, or both, (2) give examples of actual diseases, (3) examine apparent counter-examples, and (4) offer counterexamples to the contending proposals for definitions.

There are ample examples in literature of this dance of searching for necessary and sufficient conditions (see chapter 1). To illustrate, Wakefield's (1992, 1997b, 2007, 2011, 2014; 2019, 2020) work throughout the years of defending the harmful dysfunction analysis attests to the persistence and the longevity of the debate characterized by the search for necessary and sufficient conditions. A prime and recent example is Muckler and Taylor's (2020) attack on the harmful dysfunction view which they instigated by offering examples that they think are dysfunctions but are

for the most part harmless, and thus counterexamples to the Wakefield's theory.[11] They cite three examples – mild mononucleosis, cowpox that prevents smallpox and mild perceptual deficits like anosmia. Wakefield and Conrad's (2020) paper is a reply in which they show that these cases are indeed harmful and that the theory (arguably) remains unwavering against their attacks.

Now, conceptual analysis has been significantly criticized and objected to, both generally, and as we will see in the following, specifically, in the discussion on mental disorders. One common objection to conceptual analysis is that philosophy is "supposed to be continuous with science and that philosophical theories are to be defended on largely explanatory grounds, not on the basis of a priori arguments that appeal to intuition." (Margolis & Laurence, 2021, sec. 5.2 Objections to conceptual analysis). According to this view, contemporary philosophy should seriously take into account scientific and empirical evidence in order to support and explain certain claims and phenomena. These should take precedence over a priori intuitions about concepts. This is a general shot directed at conceptual analysis but it rests on the idea that at least some of the philosophical methodology is outdated and not keeping up with the scientific state-of-the-art. Dominic Murphy's project addresses at length this idea of empirical inadequacy and the problematic methodology of conceptual analysis in the discussion on mental disorders. Murphy (2006, p. 22) states that there are two ways to criticize conceptual analysis: one is that conceptual analysis is "poor analysis", the other is that it is "empirically irrelevant". He explains as poor analyses those which "fail to respect our intuitions, are open to counterexamples" (Murphy, 2006, p. 22). He deems conceptual analysis as empirically irrelevant in that it does not tell us what the concept actually is, i.e. what mental disorders are, what they consists of, but

[11] Recall that Wakefield is a proponent of the harmful dysfunction view meaning that in order for a condition to qualify as mental disorder it has to be a dysfunction that is also harmful to the person.

rather how we think about the concept, i.e. what we *think* mental disorders are. He elaborates on this later on in discussing Wakefield. His main criticism directed at Wakefield concerning conceptual analysis is that Wakefield "claims to be analyzing the scientific concept of "mental disorder"", however, he gives priority to folk intuitions over scientific and clinical concepts and empirical findings (Murphy, 2006, p. 50). According to Wakefield, if a concept revealed in our analysis corresponds to the scientific and clinical instances, but is at odds with folk intuitions, then the clinical concept is inadequate. While this might seem strange at first, Murphy contends that "this idea, that science discovers the empirical application of our pretheoretical folk concepts, is an old corollary of conceptual analysis" (2006, p. 50). Murphy (2006, p. 52) continues:

> Wakefield is criticizing the scientific, theoretical picture of mental disorder by an appeal to intuitions. This theme comes up time and time again in Wakefield's writings, as when he says that evidence for his analysis is not DSM criteria but our intuitions about disorder (Wakefield 2000).

Furthermore, Murphy (2006, p. 52) concludes that "Folk psychology's authority over the psychiatric concept of mental disorder is widely accepted by conceptual analysts.". However, Murphy ardently opposes such a view and such an analysis of mental disorder. One can interpret a bulk of Murphy's work as a testament to the idea that scientific findings, science, should have the priority in considering the concept of mental disorders rather than folk intuitions. His problem with conceptual analysis is, thus, that the way it was approached in the traditional debate undermined the role of science and scientific investigation into what mental disorders amount to. Rather, the concept was held hostage to folk intuitions, which Murphy thinks should be broken away with.

Another objection to conceptual analysis is directed at the classical view of concepts that the analysis usually relies on, which is also somewhat outdated and an

"assumption that is increasingly difficult to maintain." (Margolis & Laurence, 2021, sec. 5. 2 Objections to conceptual analysis). I will address this problematic nature of the classical theory of concepts at length in the next section.

Further objections are directed at the role of intuitions in conceptual analysis: "the intuitions that philosophers routinely rely upon may not be shared. Anyone who teaches philosophy certainly knows that half the time students have the "wrong intuitions"" (Margolis & Laurence, 2021, p. 5.2 Objections to conceptual analysis). Intuitions not only differ amongst philosophers, or students, but also with respect to people of other cultures. Concerning the discussion on mental disorder, the intuitions that are taken into account differ among those considered experts themselves, among experts and layfolk and all the combinations between various stakeholder groups in the matter. We will see all of these arguments put forth by various authors connected to the traditional discussion on mental disorder in the following sections.

In light of the criticism directed at the traditional debate, many authors are turning to other methods besides conceptual analysis. They are turning to normative analysis of the concept of mental disorder. Normative projects not only ask what the concept entails, what it is, but it also asks what the concept *should be*. I will use this dichotomy between descriptive and normative at this point just as a foreshadowing since I will focus on these in the chapters to follow. For now, it is enough to know that the method of conceptual analysis is descriptive meaning it aspires to capture how the concept is, what is the concept that might be shared by the linguistic community, while normative asks what the concept should be.

Granted, the method of conceptual analysis itself might not be the only thing to blame when it comes to the poor outcomes of the traditional debate. Pretheoretically, there may not be something wrong with conceptual analysis *per se*. It is a method, and like any other its utility and appropriateness depend on the context and other characteristics of the concept. Some of the things that are to be considered

might concern the nature of the concept we are trying to analyse, the goals and aims of our analysis, the scope of the concept we are taking into account, intension and the extension of the concept and so on. It may not be the method, but the way we are using it and with what goals in mind. When the other factors and conditions are not clearly articulated, as we will see in the following, it is no wonder that conceptual analysis gets the short end of the stick.

There are also some indications that conceptual analysis itself is not the best tool to be used even if we address all of the parameters of our conceptual exploration. As I have already mentioned, conceptual analysis is a descriptive project and it can show us how the concept is used. However, it does not give us answers to how the concept *should be* used and how to configure the concept to fit specific purposes. Some consider that to be a serious limitation concerning the psychiatric context in which the notion is immersed. The 'should' questions have gained significant traction in the recent years as the philosophers are becoming increasingly uneasy about the states of discussions based on conceptual analysis:

> The critique of traditional philosophical analysis has also generated proposals for various types of normative, revisionary projects. Instead of asking which cases are or are not covered by a concept, they ask how a concept should be modified or what new concept should be adopted in its place, given the practical context in which the concept is used—for example, given the goal of promoting social equality (Haslanger 2012, Cappelen 2018). (Margolis & Laurence, 2021, sec. 5.2 Objections to conceptual analysis)

My project is at least in part normative and revisionary which will become apparent in the remainder of the work. Because of this, I would need a bit of

methodological renovations on how to approach harm in the context of thinking about psychiatry and mental disorders since the said method of conceptual analysis does not seem to suffice. These methodological ideas will be a leitmotif of the chapters to follow.

While the primary objective of my thesis is exploration of the notion of harm in psychiatry I argue that we can take valuable lessons from the recent attempts at recalibrating the concept of mental disorder and translating them into thinking about harm. Of course, the connection between harm and mental disorder is not completely accidental or arbitrary as the concepts of mental disorder and harm are immersed in the same psychiatric context and harm is often viewed as a subset of mental disorder (though in the next chapter we will see how that does not have to be the case.).

However, what gives way to the methodological borrowing in investigating these two concepts is that the lessons that I draw out here are largely methodological and do not address so much the content of the concepts as the method of their exploration. As we have seen, the problematic elements and methods, numerous objections, apply to conceptual analysis in general, as well as to the debate on mental disorder in particular. I think that these methodological novelties and creations are a breath of fresh air in contemporary philosophy as such, not only in philosophy of psychiatry. I discuss and apply them to the debate on mental disorder as it is central in the philosophy of psychiatry and is tightly connected to thinking about harm in psychiatry. However, these ideas can be just as well articulated more generally and be applied to various other concepts that we wish to shed light on. Furthermore, this chapter provides me with an opportunity to illustrate and explain these issues on the example of mental disorders before I start applying them to the concept of harm and doing a work that is a bit more technical which will arise at points in the chapters that follow.

Let us return now to the various instances in the traditional debate that are

problematic. In the following I address these different aspects that contribute to the confusion of the traditional debate and its lack of success.

### *3.2.2 The nature of the concepts*

As I have previously mentioned, the method of conceptual analysis most often presupposes the classical theory of concepts. Some even argue that one of the appeals of the classical view of concepts is that it neatly fits with the method of conceptual analysis (Margolis & Laurence, 2021). Thus, the method of conceptual analysis and the classical view of concept seem to mutually support each other. Since both are upheld and esteemed throughout the history philosophy, their position as a status quo and their hegemony in most of the contemporary philosophical enterprises is not surprising.

The classical theory of concepts states that "a lexical concept C has definitional structure in that it is composed of simpler concepts that express necessary and sufficient conditions for falling under C" (Margolis & Laurence, 2021, sec. 2.1 The classical theory). This means that a concept can be broken into simpler elements which are considered to be necessary and sufficient conditions for the application of the concept. For example, let us take the concept 'daughter' to be a female child or offspring "in relation to their parents" (*Daughter*, 2022). We see that the concept is broken down into an assortment of more general concepts which combined specify the precise meaning of the concept 'daughter'. These in turn make the necessary and sufficient conditions for the application of the concept, meaning that if we were to call someone a daughter, they would have to be female, and be the female child in relation to their parents.

As I have mentioned, this way of thinking about concepts and conceptual analysis persisted throughout the discussion on mental disorders. It seems that one of the presuppositions of the traditional debate was that the concept of mental disorder does and can have the definitional structure upheld by the classical theory

of concepts. We have seen this in a previous section where I have indicated how the discussion was centered around the search for necessary and sufficient conditions through conceptual analysis.

In recent decades, the classical view of concepts has been contested both generally in contemporary philosophy as well as specifically in the discussion on mental disorders. All things considered, the adoption of the classical theory of concepts has a "poor track record" meaning there aren't many successful definitions that have arisen as a result (Margolis & Laurence, 2021). As Margolis and Laurence write:

> there are too few examples of successful definitional analyses, and certainly none that are uncontroversial (Wittgenstein 1953/1958, Fodor 1981). (…) The huge literature on the analysis of knowledge is representative of the state of things. (Margolis & Laurence, 2021, sec. 2.1 The classical theory)

In the antiquity, knowledge has been defined as a justified, true belief. Almost two thousand years later in 1963, Edmund Gettier in his famous paper challenged that theory by offering counterexamples to it. Thousands of papers later on the subject of knowledge and we still do not have a satisfying definition of it. This goes to show the disheartening lack of results as well as the barrenness of the classical view of concepts and conceptual analysis to produce satisfying definitions. This is in line with some of the comments that various authors make with respect to the discussion on mental disorders (Schwartz, 2007a).

Besides the unfavorable track record of such discussion, there are several renditions of the nature of the concept of mental disorder which might be at fault when it comes to lack of success of the traditional debate.

One such author is Rachel Cooper who starts by noticing that the notion of mental disorder is 'on the move'. By this she suggests that the meaning and reference of the concept tends to be shifting (R. V. Cooper, 2020, p. 148), rather than being fixed. Cooper states that conceptual change, where concepts shift in meaning is not an uncommon phenomenon, as observed by authors such as Cappelen (2018), Ludlow (2014) and LaPorte (2009). Reasons for conceptual change are various stemming from the use of some concept over time, which Cappelan (2018) calls semantic drift, because of novel empirical findings (LaPorte, 2009) or of some societal changes and pressures, such as political. For example, we can observe this with regards to the concept of marriage. With the rise of LGBTIQ rights, there is a pressure towards the shift in meaning of 'marriage' that would no longer only signify a union between a man and a woman, in places fortunate to have this shift occur.

Not all concepts are equally changeable, with some being especially prone to conceptual change. Cooper (2020) calls these human kinds, a notion she adopted from Ian Hacking (1996). Human kinds are concepts which "pick out classes of people studied by the human sciences, such as 'woman', 'black', 'long-term unemployed', 'introverted', and of course, 'disordered'." (R. V. Cooper, 2020, p. 149). These kinds are viable to change because of social movements and lobbying groups that have a stake in the matter in which they are employed and who politically and socially push towards shifts in meaning. For example, in the previous example of marriage, conservatives usually try to 'protect' the traditional definition which includes only a union between a man and a woman, while LGBTIQ activists advocate for widening of the definition to include the recognition of same-sex relationships. In this tug-of-war a concept may fall to either of the side to either keep its original meaning or be widened to include same sex marriage.

According to Cooper, mental disorder is this type of concept, a human kind term, and as such, its elusive gist is hard to capture as it is ever-changing due to

various social factors (2020, p. 149). Consequently, a project of defining a concept of disorder through conceptual analysis by seeking to determine necessary and sufficient conditions of such concepts has bleak prospects., which is how some authors describe the current state of the discussion (e.g. see Schwartz (2014)).

Schwartz (2007a) offers another hypothesis concerning the nature and the properties of the concept of mental disorder. He argues that the traditional discussion assumes the classical view of concepts where "concepts are represented by short lists of individually necessary and jointly sufficient conditions" (Schwartz, 2007a, p. 56). He lays out psychological findings on concepts that go against such view, showing that "objects are usually classified based on characteristics that are not strictly necessary", rather, the speakers classify objects in a "matter of degree, rather than an all-or-nothing affair" (Schwartz, 2007a, p. 56). For example, in the category of fruits, people usually name apples, oranges, bananas rather than tomato (this applies to Europe mostly) or passionfruit, as the former are more prototypical than latter in speaker's conceptual schema. In light of these findings, Schwartz reports that psychologists are more partial to the Wittgenstein's 'family resemblance' view of concepts, rather than the classical view.

Dominic Murphy (2006, p. 21) questions the nature of the concept of mental disorder as well. He first establishes that both strands, normative and naturalist, or as he calls them objectivist and constructivist "apply to psychiatry the regulative principle that conceptual analysis establishes what counts as the subject matter of an inquiry and the constraints under which the inquiry should proceed"(Murphy, 2006, p. 21). He dubs the adherence to this regulative principle, i.e. to use of conceptual analysis, the 'orthodox' program (2006, p. 21), what I refer to as the 'traditional' debate on mental disorders.

When it comes to the nature of the concept that the traditional debate explores, Murphy (2006, p. 21) advances a similar line to Schwartz's I have outlined

earlier. He posits that "there is no stable concept there to be analyzed, since our folk psychopathology, broadly understood, is not coherent enough to be subjected to regimentation"(Murphy, 2006, p. 22). This is evocative of a more general skepticism when it comes to concepts involved in conceptual analysis. Further, he objects that "the mind is a set of capacities rather than a clearly defined entity, and as such it can become disordered in very many ways"(Murphy, 2006, p. 22). Murphy advances this line of thinking further in his book arguing that it does not have to be the case that our folk psychological concepts that we attach to the mind corresponds to the mental processes that occur in the brain. He concludes that "this heterogeneity makes it very improbable that a simple concept, rather than a loosely united family of ideas, can serve as a model for our understanding of psychopathology." (Murphy, 2006, p. 22).

I think each of these points is important and they capture some of the important issues. Doing conceptual analysis on a concept that seems to be 'on the move' as Cooper suggests does not look very promising.

One could rightfully ask what the purpose of such analysis would be. If we take the idea of concepts as human kinds as Cooper characterizes the notion of mental disorder, then by doing conceptual analysis we are capturing various views about the concept in different points in time. Bearing in mind that the conceptual analysis is a descriptive effort, this seems to be like a survey in the sociology of language. In other words, we are essentially capturing what various stakeholders in the debate *think* that mental disorder represents at various instances in time. This in itself might still fail to grasp what the concept really consists in, if we assume that concepts are 'out there' waiting to be discovered. It fails as well if we think about concepts as being not 'out there' but confined to the minds of the language users and as such having some merit and place in the setup of how society in itself works. I would still fail to see how that would be helpful in philosophical terms since it would just tell us how we think about the concept but in contrast to the previous point our concept would then be even

more arbitrary. The project would presumably be dedicated to capturing all those different intuitions from stakeholders, but it would then just depend on the intuitions at a point in time about the concept, that seem to change as Cooper says.

The analysis also fails if we assume that concepts are things that can be constructed to procure certain benefits of using them as that would be a normative project that asks what concept of mental disorder *should* be. Conceptual analysis in its classical form cannot answer to the '*should' questions* but only to '*what is' questions*.

Thus, the use of conceptual analysis in the discussion on mental disorder falls short with regards to its philosophical merit. If the concepts are such that they do not possess a 'discoverable' essence but are just a compendium of people's ideas about the concept in a point in time, then it can be whatever the people think. As Margolis and Laurence state "if the definitions aren't there to be discovered, this would seem to put in jeopardy a venerable view of what philosophy is and how philosophical investigations ought to proceed" (Margolis & Laurence, 2021, sec. 2.1 The classical theory).

However, it is my impression from the debate that there is a difference between the concept and the discussion on mental disorders and some other debates in philosophy which rely on conceptual analysis. The debate on mental disorder seems to me to possess at all times the implicit assumption that the concept has a particular use and role in social practices. Its importance is often stated in terms of where the concept plays various functions – either in medicine, law, social work etc. As I have also stated in the 1st chapter, and as many authors have done before me, the concepts seem to matter in contexts and for concrete purposes. From this we can infer that the concept is at least in part judged with respect to its use in social practices and how well it fits into these disparate roles - medical first and foremost, but legal, ethical, therapeutic as well.

Further, I think Schwartz along the similar lines as Cooper, introduces the idea that we do not have to follow the classical theory of concepts, but we can subscribe to other theories like prototype theory that has standing in empirical, psychological evidence about concepts. Murphy advances similar criticism to that of Schwartz and Cooper about the instability of the concept.

### *3.2.3 Whose intuitions? Which goals?*

Another problem of defining the notion of mental disorder through conceptual analysis is the question of who are the 'competent users', whose intuitions are we trying to capture? The goals of various authors seem to differ which is why one could get the impression of the main authors talking past each other. Lemoine (2013) points to this disparity in goals of different authors. His efforts when he does seem to be directed more at outlining and systematizing the main contenders of the debate – Boorse, Nordenfelt and Wakefield, rather than criticizing it. In the following I address these problems building on Lemoine's (2013) analysis.

Boorse's aim is to configure a theoretical notion of mental disorder, which differs from the practical notion. The account he develops aims at "the pathologist's concept, not a clinician's and still less at any social or legal category" (Boorse, 1997, p. 11) of disease. His goal is to analyse mental disorder in analogy to the view of disorder "as understood by traditional physiological medicine" (Boorse, 1977, p. 543).

Conversely, Wakefield (2007, p. 150) considers one of the central goals to "clarify and reveal the degree of legitimacy in psychiatry's claims to be a truly medical discipline". The approach most relevant for that goal is "conceptual analysis of the existing meaning of "disorder" as it is generally understood in medicine and society in general" (Wakefield, 2007, p. 150). In another place, he says that: "The concept is largely shared by professionals and lay public (Campbell, Scadding, & Roberts, 1979) and is the basis for the attempt DSM – III – R to construct universally acceptable atheoretical diagnostic criteria" (Wakefield, 1992, p. 374). Wakefield thus takes into

account intuitions from medical experts, both in research and clinicians, as well as the intuitions and ideas about mental disorder of lay people, basically competent users of language. The latter are also known as *folk psychological* intuitions about mental disorders (for a discussion see Murphy (2006)).

We can observe this variety in goals in other prominent authors in the debate as well, as Lemoine (2013) points out. Nordenfelt (1993, p. 17) uses a "framework of health concepts" that include disease, health, illness, sickness, defect, injury, impairment and disability. According to him, "a fruitful discussion of the concept of disease presupposes that this whole network of concepts in particular the concept of health is considered."(Nordenfelt, 1993, p. 17). He lays out different contexts in which these concepts are being used. He maintains that in addition to scientific contexts,

> these concepts basically belong to the discourse of ordinary life. They are frequently used in our day-to-day affairs, when we report on our bodily and mental states and the stare of our fellow human beings. Their frequent use is understandable. (Nordenfelt, 1993, p. 17)

In addition, Nordenfelt lays out important contexts in which the concept of disorder is used – health and disease as the focus of the philosophy of welfare, in health care settings, in the clinical context, in the clinical sciences and in the context of moral responsibility. Nordenfelt (1993, p. 24) lends that in all these various contexts the concepts might not be used coherently nor rationally. However, we "have to know the meaning of the concepts in these context" to establish that, for which he prescribes conceptual analysis.

We see that three of the most prominent authors in the debate, who among others also debate with each other, may have different goals. Now, while this does not necessarily mean that these concepts do not overlap, as both lay pathologists, lay

people, use of disorder in legal and moral contexts could, in theory, talk about the same concept. However, this seems highly unlikely due to the diverse usage of the concept and the various parts that get highlighted in different contexts. For example, in legal sphere the concept of mental disorder is usually employed in determining legal responsibility while in psychiatry is used to identify and treat pathological state in individuals. Since the concept is used in numerous areas and serves different purposes it would be a stretch to say that in these myriads of cases we are dealing with essentially the same concepts. Even more so, these concepts may even be incommensurable up to a point. This would mean that because of their roles and presumable intentions as well as extensions, the concepts are not really comparable. If this were the case, it would make communication between social practices that utilize the concept that much harder. It would also render the aspiration of finding the comprehensive and one-size-fits-all definition of mental disorder ineffective. As Lemoine (2013, p. 312) asserts at one point about the authors of the traditional debate: “the apparent variety of these goals does not prevent them from arguing against each other”. Maybe this can be generalised to the debate as a whole and it might be one of the reasons why we have not reached a satisfactory account of disorder. It might also illuminate the reasons why the efforts at defining the notion of mental disorders have not come to fruition.

Connected to the previous point, in addition to the different intension of the concept of mental disorder of these authors, the extension of the concept of mental disorder differs among participants in the debate as well.[12] They do not take the same set of cases for the notion to refer to, as Lemoine (2013, p. 314) points out. Boorse (1977, p. 551) uses the Standard Nomenclature of Diseases and Operations published

[12] Intension of the concept is its internal reference, meaning that it captures, for example, the intension of the notion 'bachelor' is 'an unmarried man'. Extension of the concept are the cases in the world which are captured by the concept, i.e. all men to which the term bachelor applies.

by the American Medical Association which he considers to be “invaluable as a compendium of thousands of test cases for an analysis of disease”. However, he does not deem it as a ‘completely authoritative’ as the “evidence must be combined with the usage of 'disease' and 'health' in the discursive context of medical textbooks and research papers” (Boorse, 1977, p. 551). Thus, he remains encroached in, what he would refer to as the theoretical medicine and its rendition of mental disorder.

Wakefield, conversely, uses a more vaguely defined set of cases taking into account more of a general consensus on what counts as a disorder (Lemoine, 2013, p. 315). Recall, Wakefield considers a variety of intuitions ranging from medical professionals to lay people about what counts as a disorder. As Lemoine suggests, this might make it harder for Wakefield to distinguish controversial from uncontroversial cases. Additionally, it makes the set of cases that are the extension of his concept of mental disorder dubious.

Nordenfelt (1995, p. 1), contrary to Boorse, rejects the view of disorder from the standpoint of medical experts as authoritative:

> It is important to emphasize that medicine has no monopoly on the concept of health or related concepts. Health, impairment, disease and disability are concepts which are well embedded in ordinary thinking and which have a long non-scientific tradition. There is an ordinary concept of health which can be used by the layman with the same accuracy as he can use most other central concepts characterizing man, for instance concepts of morality, emotional concepts, or concepts of excellence.

Nordenfelt, thus, gives at least the same amount of weight (if not more) to the lay conceptions of mental disorder as he does to those by medical experts. This however does not give us a conclusive result as to the extension of Nordenfelt’s

concept.

Hofmann points this problem with different extensions of the concept as well saying that "the various disease concepts do not always have the same extension (…) Sometimes the concept refers to the theories of disease, while in other instances the ascription of disease is the issue, and sometimes they are intermingled" (Hofmann, 2001, p. 218 citing Räikkä, 1996).

It may be the case that the extension of the term disorder overlaps to such extent that the differences between them are not contestable. However, we do not know that and the idea that they are all taking the same set of cases into account is highly improbable. The problem worth emphasizing is that the set of cases is not clearly defined among authors. Conceptual analysis of the concept whose intension is unclear and differs among authors, and whose extension differs among authors as well, meaning there are no explicitly defined cases that are taken into account, poses considerable difficulties in coming up with solutions. As Nordby claims: "For if participants in the debate disagree about what states should fall under the definition, then they will necessarily end up with different definition." (2016, p. 174)

### *3.2.4 One concept or many?*

Furthermore, it is not clear whether we are looking at one concept of mental disorder or many. Granted, some authors in the discussion, like Boorse, are looking for a single concept of mental disorder to do specific work in theoretical medicine. Other authors, like Nordenfelt, might be looking for a more universal concept that would capture the lay understandings of the disorder, medical practice and law and other areas in which the concept is applied. However, authors like Nordby (2016, p. 178) argue that we can only aim for contextual definitions of mental disorder as "It is only possible to formulate pluralistic definitions that capture different sets of assumptions about the use of the concepts.". He argues that it is integral to identify and explicate conceptual assumptions of different areas in which the concept is used and concludes concerning

the health concepts that “it is more promising to connect them to special areas of discourse” (Nordby, 2016, p. 179).

In the similar vein, another author, Juha Räikkä, brings into question whether the debate is about one and the same concept. It seems that the authors, on both naturalist and normativist side, “are talking about the same sense of “disease”, namely the *medical concept of disease*”(Räikkä, 1996, p. 353, original emphasis). However, the problem arises because the stakeholders in the debate are trying to ascribe numerous characteristics and properties to the medical concept, especially those pertaining to the ‘social’ rather than the ‘medical’ sphere. According to Räikkä, this is a mistake because “contrary to common wisdom, the concept of disease used in discussing social questions is not the medical concept of disease, but what might be called the *social concept of disease”* (Räikkä, 1996, p. 355, original emphasis).

What is his rationale for these claims? He postulates that regardless of the position one takes in the debate, whether it be naturalist or normativist “to label a condition a disease in the medical sense is not, *merely by that judgement*, to indicate any norms or directions for society” (Räikkä, 1996, p. 357). He divides the medical and the social sphere saying that in the use of the social concept of disorder we “clearly indicate norms and directions for society”(Räikkä, 1996, p. 357) while also arguing that these are not derived from the medical concept of mental disorder. Rather, he argues that:

> When a given condition is called a disease in the social sense of the word, there is a concomitant judgment that a certain act or practice is justified or obligatory. The social concept of disease fits our everyday discussions concerning diseases. *By definition*, it is a concept according to which disease-attributions are norms and directions concerning the treatment of those who exhibit conditions called diseases. (Räikkä, 1996,

p. 357)

Thus, according to Räikkä we can adduce that the problems in the debate stem from the conflation of the medical and the social concept of mental disorder. Normativism in the debate of the medical concept of mental disorder does not designate judgments about treatment and the social reactions to certain disorders. Rather, that is the function of the social concept. Because the social concept of disorder is mistakenly reduced to the medical concept the debate relishes in unsolvable issues pertaining to the different roles and functions ascribed to the medical concept. As he argues, however we define the medical concept from the normative side, it does not contain information about the social response to certain conditions – whether we should treat, how and in what way we should 'ease' the social integration and response to these individuals.

Räikkä's analysis has merit. We can suppose that the division between medical and social concept could be helpful in solving some of the perplexities of the debate. However, it does necessitate a thorough rearrangement of the traditional debate or branching out into defining two concepts instead of one.

It is not clear, though, whether we can assume such a sharp divide between the medical and the social concept. It seems that in the psychiatric context, the medical concept of disorder and particular diagnoses of disorder do carry with them some implicit information about patient care. They may not be tied directly to it but I do not think it is clear that from the medical concept we cannot derive information about treatment and social care. After all, this is how the concept of disorder is used now. The diagnosis from the psychiatric manuals is used in various contexts and contains information about the condition, including the symptomatology and the characteristic manifestation of the conditions. It certainly might be beneficial to explicate these, even to synthesise these into a social concept

different from the medical one.

Bearing in mind the other inconsistencies in the debate that I have covered so far, I think the divide between the medical and the social concept would entail careful reconfiguration of many of the issues that I have already mentioned as being murky in the debate. However, it is not clear that by solving and specifying all of the issues and problems with the debate we would end up with two separate concepts. It is helpful to know that dividing the concept up into two is an option that can be used, come rain or shine.

### 3.3 Conclusion

The focus of this chapter was to outline and systematize an array of reactions towards the traditional method of conceptual analysis that is becoming more and more prominent in contemporary philosophy. I have covered the objections directed to the method of conceptual analysis with respect to the discussion on mental disorder, the nature of the concepts involved in the discussion, the question of intuitions and goals that we want to achieve with our concepts as well as with the number of concepts that we are discussing. Most of these are connected to the debate on mental disorder. Since harm resides in this discussion and is connected to the methodological and theoretical customs of the debate, exploring the harm criterion necessitates answering much of these methodological questions as well.

However, in the next chapter we will see how harm can theoretically be a stand-alone criterion in psychiatry. There is one line of thought that can be characterized as most severe in its disapproval of the debate and it offers the most controversial solution. This is eliminativists with regards to the notion of mental disorder. In the following section I address and outline their positions while arguing that even if we adopt eliminativism, harm still has a rightful place in psychiatry.

# 4. Eliminativism

## 4.1 Introduction

In the previous chapter I have outlined the problems that many authors put forward with regard to philosophical methodology in general, and with the traditional debate on mental disorder specifically. As a result, some authors in contemporary philosophy of psychiatry propose the dissolution of the traditional debate on mental disorders. Even more so, they argue that we should do away with the concept of mental disorder entirely. This is the position called eliminativism with regards to the notion of mental disorder. I dedicate this chapter to outlining this somewhat controversial position. However, rather than arguing for or against it I argue that even if we eliminate the notion of mental disorder, harm still has its rightful place in psychiatry. I first present the eliminativist positions in the debate defended by Hesslow, Ereshefsky and Bortolotti. I then explore how these eliminativist ideas relate to the notion of harm. I argue that even if we adopt eliminativism, harm remains a relevant criterion inside the psychiatric practice. Next, I discuss how harm relates to the notion of benefit as many see these two concepts as complete opposites of each other. I show that this is not necessarily so and that in medicine it makes more sense to focus on reducing harm than procuring benefits. Finally, I reconsider the idea of eliminating the notion of mental disorder and offer reasons why that might not be the soundest solution.

## 4.2 Eliminating the Notion of Mental Disorders

Eliminativists with regards to the concept of disorder think that medicine does not

need the concept of disorder and that it can afford to do away with it. There are several authors that advocate eliminativism, most notably Germund Hesslow (1993), Marc Ereshefsky (2009) and Lisa Bortolotti (2020).

Hesslow's paper from 1993 entitled "Do we need a concept of disease?" is a seminal work arguing for eliminativism. Hesslow starts with an analogy. He asks us to imagine a car owner going to the mechanic because the car does not accelerate as fast as he would want it to. The mechanic asserts that the car is not defective, that there is nothing wrong with it, but rather that the engine is just adjusted in a way which is not conducive to the acceleration that the car owner wants. This sets off a polemic between the car owner and the mechanic about whether the car is defective or not. Hesslow (1993) claims that it would be useless for the car owner and the mechanic to engage in such a discussion. Rather, the car owner should tell the mechanic that they want the car set up so that it accelerates faster and the mechanic has the reason to comply with his demand regardless of whether the car is defective (Hesslow, 1993; Stegenga, 2018). Analogously, Hesslow (1993) maintains that the concept of mental disorder is useless:

> The health/disease question is irrelevant - we never really need to know whether someone has a disease or not, and consequently, we do not need a definition of 'disease'. (Hesslow, 1993, p. 2)

Rather, we treat according to grievances directed at certain states regardless of the disorder status. In this sense, Hesslow thinks the notion of disorder to be superfluous. Sure, it might play some role in everyday discourse but in sophisticated scientific and medical contexts we do not really need to know whether someone is 'diseased'.

There is doubt whether the analogy Hesslow presents is convincing. As Jacob Stegenga (2018) notes, if a car is advertised and guaranteed by the manufacturer to

have the capacity to accelerate in a particular way, then it seems that the car is defective with regards to that standard. Thus, as Stegenga (2018) emphasizes, the car owner has a right to ask the mechanic to repair the car according to that standard, and have the car manufacturer pay for the repair. The same would go for mentally disordered who in the case of substandard functioning and performance are granted rights to treatment in light of their disorder status, as well as additional social and legal rights that stem from it. Thus, he contends that the notion of disorder has important economic and social consequences (Stegenga, 2018).

Hesslow (1993), however, dedicates the remainder of his article to debunk all of the reasons why defining the notion of disorder is presumed to be useful in scientific inquiry and in social and moral contexts. Hesslow (1993) describes his project as both descriptive and normative. It is descriptive in that the notion of disorder does not play that important of a role as it is, and it is normative in that where the notion of disorder does play a role, Hesslow (1993) argues that it should be reduced.

First, he covers the cases of scientific inquiry where conceptual analysis is useful. To name a few reasons, some concepts are integral to scientific theories and as such have important "intellectual and practical" consequences as they determine the content of a theory (Hesslow, 1993). Additionally, definitions of concepts such as time, space, and motion can be important for understanding of a theory, for example, theory of relativity (Hesslow, 1993). He contends that none of the cases where defining the concept is important with regards to scientific inquiry applies to the concept of mental disorder. He concludes:

> There is no biomedical theory in which disease appears as a theoretical entity and there are no laws or generalizations linking disease to other important variables. Therefore, there is no need for an analysis that makes

> a theory more determinate or understandable, and there is no need for operational definitions. (Hesslow, 1993, p. 5)

Next, Hesslow (1993) addresses the idea that there might be some social or moral reasons behind defining the notion of disorder, of which he covers four of them: 1) disease as a grounds for medical treatment, 2) uses of the notion of disease in cases of medical insurance, 3) disease as a grounds for special rights (sickness benefits, working ability), 4) mental illness and responsibility.

Here I will address 2),3), and 4) to outline the eliminativist view and will return to the 1) later when addressing eliminativism in connection to harm.

In relation to the notion of disorder and medical insurance Hesslow (1993) argues that the notion of disorder does not do much work there and even if it did, it does not need to correspond to the disorder status. For example, in some cases insurance might not cover conditions which are justifiably considered disorders but that the person has brought onto themselves, like skiing accidents or smoking related diseases. When it comes to disorder status granting some persons special rights, for example those that are work related, the thing is also not determined by the disorder status per se, rather by the nature of the condition and the nature of the work the person does.

The question of criminal responsibility and mental disorder is also not determined by the disorder status in itself but by "the inability of the individual to be influenced by the consequences of her actions" (Hesslow, 1993, p. 10). In other words, it is determined by the individual's ability to assess the wrongfulness of their actions and the responsibility one can assume over them. Lisa Bortolotti (2020) argues in a similar vein that the cases of criminal responsibility should be decided in terms of whether the person's agency is compromised. She also notes that the same person with a diagnosis of mental disorder, for example schizophrenia, can be morally and

legally excused in some circumstances while being responsible in other circumstances as their agency can be affected differently at various points in time (Bortolotti, 2020).

Thus, Hesslow (1993) and Bortolotti (2020) conclude that even though the concept of disorder might be useful in some instances and in everyday discourse, its importance for medicine as well as society is marginal at best.

The eliminativist view is somewhat controversial and it faces objections. The dispute is generally centered around the extent to which the notion of mental disorder plays important scientific, social or moral roles.

Lennard Nordenfelt (1993) defends the notion of disorder against Hesslow's attacks in the following way. First of all, he widens the scope of what he is addressing from just the notion of disorder to what he calls the "network of health concepts" which include health, illness, injury, impairment, disability and defect (Nordenfelt, 1993, p. 17). He underlines that these concepts are not exclusively medical nor entirely scientific but that they have a social and a political role, a point that is quite salient throughout his paper (Nordenfelt, 1993). Importantly, he points out three different reasons the concept of disorder is worth exploring philosophically:

> "(i) the health-concepts being of primary concern in ordinary life and political life,
> (ii) their having a great deal of complexity, and,
> (iii) their being the object of considerable controversy" (Nordenfelt, 1993).

He argues that these health-concepts are widespread because they are a matter of welfare, of how a person's life is going for them, and that has an important role in social, as well as political discourse. Additionally, the health-concepts do have a certain complexity, and they are a matter of controversy, for example, the cases of homosexuality or alcoholism (Nordenfelt, 1993).

Nordenfelt (1993) contends that Hesslow undermines the sociological role of

medicine emphasizing that judgements ”about health and disease can have important social and legal consequences.” (Nordenfelt, 1993, p. 20). One monumental example of that is the case of demedicalisation of homosexuality that is quite present in the literature. In brief, homosexuality was first considered as something immoral and illegal, which stemmed from both religious and legal authorities in the Western countries (Nordenfelt, 1993). Then, this changed once homosexuality was recognized as a mental disorder, as something pathological, meaning it ceased to be a criminal offence and became a mental disorder.[13] This however, changed in 1973 when American Psychiatric Association removed homosexuality from the Diagnostic and Statistical Manual (1980), which marked the depathologization of homosexuality (Bayer, 1987; Drescher, 2015; Nordenfelt, 1993). The journey from decriminalization to depathologization and further was no small feat, and it can be largely attributed to the activists of the Gay Liberation Front and other social initiatives (Nordenfelt, 1993).

The depathologization of homosexuality, the shift from it being a disorder to it being pronounced as a healthy variant, has had a tremendous social, political and cultural effect in terms of the recognition of gay rights. Thus, Nordenfelt (1993), pace Hesslow (1993), emphasizes that whether something is a disorder or not is certainly not a moot point.

So far, the cause of disagreement has been the degree to which the disorder status and the concept of mental disorder plays a role in various contexts, be it medical, scientific, social, or political. Hesslow (1993) maintains that it does not really play a key role in these contexts, and where it does, it should be reduced. Nordenfelt (1993) argues that it plays important social and political roles. It seems that the disagreement is a factual one and it can be adjudicated by empirically investigating the role the concept plays in various contexts. That goes for the descriptive part of

[13] It coincided for quite some time as noted in (Hooker, 1956) there was a triple stigma attached to homosexuality: it being a sin, a crime and a disease.

their arguments.

Hesslow, however, goes further and argues a normative point that the notion of disorder should not play the role that it does, where it does. He argues that in sophisticated medical contexts the notion of mental disorder is naturally replaced by some specific technical term which specifies the problem and is in the jargon of the discipline.

In line with this, what can be objected to Hesslow (1993) and Bortolotti (2020) is that while they criticize the project of defining the concept of mental disorder, and the usefulness of the concept itself, they do not offer some alternatives on what to focus instead. For it seems that the notion of disorder plays at least *some* role and that the concept has at least *some* importance although Hesslow wants to denigrate it. Some guidelines on how and what to use instead of the notion of mental disorder if we eliminate it seems like a reasonable request.

In contrast, Marc Ereshefsky, another eliminativist who criticizes the traditional discussion on disorder, provides suggestions on what the concept of mental disorder should be replaced with. Ereshefsky (2009) addresses the problems with the discussion on mental disorders between normativists, naturalists and hybrid theorists. He assesses the situation of the discussion as a stale mate, which does not seem to get resolved. Therefore, he proposes alternative criteria to the concept of mental disorder that we should pay attention to.

Ereshefsky (2009) notices that throughout the discussion on the notion of health and disease, two notions come to the forefront: *state descriptions* and *normative claims*. State descriptions are "descriptions of physiological or psychological states" that some patient is experiencing, as he gives the example of red blood cell rupturing, or the amount of calcium in patient's tissue, or some psychological processes like emotions (Ereshefsky, 2009, p. 225). They are purported

to be devoid of as much normative elements as is possible.[14] This would serve the purpose of the objective, scientific components in the discussion, usually advocated on the naturalist side of the debate on mental disorders. However, it would avoid notions such as function and dysfunction as they breed controversy in the discussion and they sometimes make normative assumptions (Ereshefsky, 2009, p. 225).

The other element, normative claims, are "explicit value judgments concerning whether we value or disvalue a physiological or psychological state." (Ereshefsky, 2009, p. 225), for example, we disvalue broken legs, ruptured aortas, and all kinds of other description states. As Ereshefsky (2009) notes, making these values explicit falls under the notion of normative claims.

Ereshefsky (2009) emphasizes several upshots of using the distinction between state descriptions and normative claims rather than trying to define the notion of mental disorder. First, this dichotomy helps to clear out some controversial medical cases like deafness, where some argue that deafness is a disorder while others argue that deafness is part of their identity and it affords them affiliation with their deaf community (Ereshefsky, 2009). According to Ereshefsky (2009), the point of disagreement is not in state descriptions, as both sides agree with the physiological properties of deafness, but in normative claims as one group values the state of deafness while the other disvalues it. Similar can be said about the movement that started to arise concerning the conditions traditionally falling under the psychiatric domain, like autism or Asperger's, where it is presumably disvalued by medical practice while the neurodiversity movement claims that autism, or at least some instances of it, should not be disvalued as it is a part of the person's identity.

Another reason Ereshefsky (2009) has for the dichotomy he proposes is that it

[14] Some authors, like Cooper (2020) and Kingma (2014) argue that there are different kinds of normativity and on different levels of scientific understanding, so that some normativity is unavoidable.

avoids the problems of using contestable notions such as normal, natural and dysfunction that arise in the discussion between naturalists and normativists. Further, he argues that it avoids some further problems which face the projects of defining health and disease (Ereshefsky, 2009).

All in all, the take-home message of Ereshefsky's account is that the discussion on the definition of mental disorder makes two considerations salient: state descriptions and normative claims. His proposal is to concentrate on these notions and reframe our thinking in such a way rather than trying to find a definition of disorder.

Now that I have outlined the main eliminativist positions in the debate, let us focus on how harm fits into this picture. As a reminder, whether to keep or to eliminate the notion of disorder is not the focus of my work. Instead, I focus on the notion of harm. Thus, investigating the move of eliminating the notion of disorder is relevant so long as we consider the way this would impact harm. In the following I address the issue of what happens to harm if we suppose eliminativism with regards to mental disorder. I argue that harm is here to stay regardless of if it finds a place inside or outside of mental disorder.

### 4.3 Eliminativism and Harm

As a reminder, one fraction of the discussion on the definition of mental disorder, which I have outlined in the first chapter, has to do with the role and the significance of harm as one of the criteria of mental disorder. While some authors advocate for the removal of harm from the definition of mental disorder (Amoretti & Lalumera, 2019; Feit, 2017) others argue for its inclusion in the definition (R. V. Cooper, 2002; Wakefield, 1992).

At face value and acquainted with the discussion on the definition of mental disorder that is presented in the first chapter, one might assume that the existence of harm and its place in psychiatry depends on whether it is a part of the definition of

mental disorder. Since in the traditional debate, harm is a normative element present in normative and hybrid theories, one might assume that there is room for harm so long as there is a definition of mental disorder, and if the definition affords itself, at least in part, to the normativist view of disorder. However, in this meta-discussion the role of harm needs additional attention and analysis. I urge us to think of harm outside of the 'box' of the traditional discussion on mental disorder. I argue that we should think about harm in terms of elements that are relevant for the psychiatric setting. I call these *the criteria of psychiatric interest.* These signify the criteria in psychiatric diagnosis, assessment, treatment, and any other area of psychiatric context which are particularly salient and that play an important role. Since the whole idea of the notion of mental disorder is brought into question in this meta-debate, I use 'criteria of psychiatric interest' to cut across the traditional debate and extract what is relevant in terms of my thesis.

I argue that harm, as a criterion of psychiatric interest can find its place in psychiatry regardless. Harm plays an important role whether it be outside of the notion of disorder, as a criterion of the notion of disorder, or standing on its own without a notion of disorder. Thus, even if we suppose eliminativism, I argue that harm has an important and recognizable role in the psychiatric setting, as one of the criteria of psychiatric interest.

Let us go back to Ereshefsky's idea that the notion of mental disorder should be eliminated and replaced with state descriptions and normative claims. I contend that harm elegantly fits into that picture. As I have described, state descriptions are an objective, matters-of-fact element while the normative element is the one that does the evaluative, value-laden work. The notion of harm could be fitted into this two-partite frame as being a normative, value-laden element. Harm already is considered a normative element in the traditional discussion on mental disorder I have outlined in the first chapter. Thus, it seems rather uncontroversial to think of it

as such in the Ereshefsky's theoretical framework.

Thus, in evaluating conditions of medical, or specifically psychiatric interest, in determining the state description we would have to ask ourselves whether this state is harmful to the person or not. Asking this question explicitly might open up new avenues and opportunities for discussion. Furthermore, the notion of harm might even help solve some of the borderline cases. Let us look at the case Ereshefsky mentions of a deaf person. While ex hypothesis the state description is agreed upon, asking whether the deafness is harmful to the person might shed some light on the issue. Of course, it depends which account of harm one adheres to and on particular properties of harm, but it drives the discussion in a helpful and constructive direction, which I show to be the case later on.

For now, it is important to recognize that harm is consistent with the Ereshefsky's flavor of eliminativism meaning that its role can be recognized even if we discard the notion of mental disorder and replace it with state descriptions and normative claims.

Let us now focus on Germund Hesslow (1993) and Lisa Bortolotti (2020) view on eliminativism and discuss how harm relates to their views. As I have already mentioned, Hesslow (1993) addresses the idea that there might be some social or moral reasons behind defining the notion of disorder which he tries to refute. Here I focus on the claim that 'disease as grounds for medical care' which is one of four claims that Hesslow seeks to dispute in his paper (see 2.1. for other three). This one is particularly important for the notion of harm. Hesslow contends that although medicine usually treats diseases, there are many cases where the medical treatment does not correspond with disease. For example, there are conditions which are reasonably considered diseases, like benign tumors, birth marks or fibromas, but are not treated because they do not pose danger or discomfort to the person (Hesslow, 1993). Conversely, there are conditions where there need not be disorder to facilitate

treatment by a physician, for example, cosmetic surgery or sex change operations (Hesslow, 1993). Thus, Hesslow concludes, although disease is a useful label for justifying medical intervention, it is not crucial. Rather, Hesslow thinks that what is crucial is "the fact that some medical intervention may be beneficial and that it is within the physician's power to help the patient." (1993, p. 7). Let us call the first of these, 'the fact that some medical intervention may be beneficial' the criteria of benefit and, the second, 'the physician's power to help the patient', the criteria of treatability. Hesslow (1993) thinks that these, benefit and treatability, are sufficient conditions for seeking medical attention.

Let us now turn to Lisa Bortolotti (2020) who discusses the issue of eliminating the notion of mental disorder similarly by providing the example of lower back pain. She argues that the symptom can be treated in various ways and the cause of it can also be versatile ranging from a slipped disc, a sciatica, or an unidentified cause (Bortolotti, 2020). She contends that there are different ways of approaching the problem and different methods of intervention, however, it is not the case that the disorder status needs to be determined beforehand, especially in the form of biological dysfunction (Bortolotti, 2020).

In general, the appropriateness of the treatment varies meaning that not all people who are treated are disordered (pregnancy), and not all who are disordered are treated (Bortolotti, 2020). This is a similar point to the one I have outlined above made by Hesslow where he emphasizes that treatability and disorder do not entirely overlap. Therefore, although medical treatment is connected to the disorder status, it seems that both Hesslow and Bortolotti maintain that it does not have to be and that not much is lost if it is not.

Bortolotti (2020) further emphasizes that not only does the view of medicine as preventing and treating diseases inappropriately correspond to the extension of the notion of disorder, but it is also too narrow and not really representative of all the

roles that medicine plays in our lives. According to Bortolotti, we are justified to seek and receive medical advice for perceived problems in our lives without any talk of the notion of disorder, for example regarding the problems like the inability to sleep, tiredness, nervousness and so on.

Applying her considerations to the cases of delusions, Bortolotti (2020) argues that harm is superior than naturalist accounts in determining what justifies medical attention: “the presence of harm does seem a sufficient reason for seeking help from medical professionals, and this applies both to delusional beliefs and other beliefs associated with experiencing harm.” (Bortolotti, 2020, p. 177)

Thus, according to Bortolotti (2020) harm is sufficient for seeking medical attention regardless of the disorder status. Hesslow, as we have seen argues that incurring a benefit and treatability of a certain condition are sufficient grounds for medical attention. These are views that are quite similar.

Specifically, harm and benefit are notions that are often juxtaposed and considered opposites, i.e., benefitting someone is the opposite of harming them and harming someone is the opposite of benefitting them. Thus, by parity of reasoning, incurring a benefit can be equated with reducing harm, and harming someone can be seen as benefit reducing.

If this were always the case, we could say Bortolotti and Hesslow are stating exactly the same thing in that the sufficient criterion for medical care is either reducing or preventing harm or incurring benefit. When these occur simultaneously, meaning that when harming someone necessarily entails reducing benefit and reducing benefit necessarily entails harming someone, we are in fact talking about the same thing. For example, breaking someone’s leg entails harming them in virtue of causing pain to them, and disabling them from walking, running and so on. This can also be interpreted as reducing, or taking from them the benefit of painlessness, and the benefit of walking, running, etc. Conversely, giving someone a pill for a headache

entails bestowing a benefit to them in taking away their pain, and reducing harm, the pain. Thus, when these are two sides of the same coin, we can say it does not really matter whether we speak of providing benefit or reducing harm. However, there is an important twist where not all benefit equals reducing harm, where these notions come apart and are not completely complementary. Because these senses come apart at some point, and imply different things, I argue we should give primacy to harm rather than benefit in the medical context. I provide examples from medicine to support my claim.

In the following I discuss the idea that there is a difference between incurring benefit and reducing harm. Considering this difference, I argue that psychiatry, and medicine more generally, works and functions more in the sense of reducing harm than incurring benefit, at least in the most relevant, nonproblematic cases. I argue that therefore it makes sense to focus and to give priority in our considerations to harm.

### 4.4 Harm or Benefit?

Harm and benefit are some of the central notions in moral philosophy. Although here we are dealing with the philosophy of medicine, there are useful insights we can adopt from this literature, which will provide us with a more nuanced understanding of both harm and benefit.

Seana Shiffrin (2012) is one of the authors who discusses the relation of harm and benefit in moral philosophy. In discussing the priority of harm, Shiffrin outlines two asymmetries of harm and benefit. The first asymmetry:

> arises between the differential reason-generating force and the reason-requiring force exerted by harms and benefits. Generally, other things being equal, harms, harming events, and opportunities to harm are more important morally than benefits, benefiting events, and opportunities to

benefit. (Shiffrin, 2012, p. 361)

In other words, harm seems to have some precedence and importance over benefit. Harm gives stronger reasons and has more capacity to provide reasons for action (Shiffrin, 2012, p. 361). As Shiffrin (2012) notes, given a harm and a benefit of equal size and significance, it seems that preventing or alleviating the harm is more important than providing a benefit of equal importance. Applied to medicine, it seems that at least prima facie the same holds. Reducing, preventing and alleviating harm is more important than bestowing a benefit of equal importance (as a reminder, under the assumptions that they do not overlap – the situation where it is either harm reduction *or* inducing a benefit).

Many ethicists argue that preventing, curing and alleviating disease is and should be the primary goal of medicine and that medicine should be defined as such (Brody & Miller, 1998; Callahan, 1996; Miller et al., 2000, 2000). For example, an international group of scholars in The Hastings Center Report, a publication which explores ethical, social, and legal aspects of medicine, recommended a list of four goals:

> "the prevention of disease and injury and promotion and maintenance of health"; "the relief of pain and suffering caused by maladies"; "the care and cure of those with a malady, and the care of those who cannot be cured"; and "the avoidance of premature death and the pursuit of a peaceful death." (Callahan, 1996, pp. 10–14)

Here we see that emphasis is on reducing harm as the goal of medicine as each of these goals can be subsumed and described as such. I think these cover most medical interventions that are undertaken.

Admittedly, a portion of medicine is dedicated to primarily bestowing benefits rather than prioritizing prevention and reduction of harm. For example, plastic

surgery for aesthetic reasons, contraception, abortions, removal of skin tags and birthmarks, sports medicine in terms of performance maximization. However, there are reasons why these are more of exceptions than the rule. First, it seems that most of the medicine is still dedicated to reducing and preventing harm even if there is a small portion of it dedicated primarily to bestowing benefits. This goes as far as some authors claiming that harm reduction rather than bestowing benefits is *the right* thing to focus on arguing that it is inappropriate for medicine to bestow benefits in absence of disease (Kass, 1975).

Second, indicative of the primacy and the special role of harm reduction is also the fact that these procedures –for example, aesthetic surgery, contraception, and abortion, are in the majority cases not covered by medical insurance but privately by patients themselves. This might reveal the implicit assumption that medical insurance covers only instances of disease (Schwartz, 2014, p. 582). For example, in Croatia functional rhinoplasty, colloquially known as a "nose-job" with a purpose of restoring or improving respiratory function of the upper airways is covered by medical insurance, while aesthetic rhinoplasty, a "nose-job" for purely aesthetic reasons is not.[15] We can thus conclude that the priority is harm alleviation, prevention and reduction rather than incurring benefits.

The second asymmetry of harm and benefit has to do with what is allowed to be done to avoid or alleviate harms versus what is allowed to provide benefits (Shiffrin, 2012, p. 363). To illustrate this, Shiffrin discusses what she calls the Rescue case. It seems that it is generally uncontroversial that one may cause some harm to the person to alleviate greater harm, for example, doing a necessary, possibly life-saving operation on an unconscious person unable to consent to it, or rescuing a drowning person even though it involves breaking their arm (Shiffrin, 2012, p. 363). Yet, when it comes to benefit it seems that the same does not hold in that generally

[15] This is anecdotal as I have a deviated septum which might require surgery.

we are not allowed to cause some lesser harm to bestow a benefit onto someone, for example, if breaking someone's arm could improve their memory or hearing so that it is supersensitive we would not be justified in doing it (Shiffrin, 2012).

One might argue that there is no such benefit that could be equated with the harm of losing a life, but it seems that if we lower the stakes the holds – we are allowed to do lesser harm in order to prevent greater harm rather than to do a lesser harm to bestow a greater benefit. If we had an option to break someone's leg to save their leg, for example, if their leg was stuck somewhere, or to break their leg to incur some benefit like being able to jump really high or fast with that leg once it heals, it seems that the former is justifiable while the latter is not.

In medicine the same holds, as can be adduced from consent in the medical context. Informed consent is the autonomous approval of the patient to participate in some medical procedure or research (Terry, 2007). Ordinarily, informed consent to treatment consists of several elements which must be facilitated by the physician for patient's consent to be valid:

> (1) disclosure of the pertinent medical facts and alternative courses of action; (2) ensuring patient capacity to understand the decision to be made; (3) ensuring patient understanding of the medical information; (4) the absence of coercion or manipulation; and (5) the ability to consent. (Terry, 2007, p. 564)

In most medical cases this procedure must be followed to ensure the protection of patient's rights and autonomy. However, there are cases where the informed consent is not possible. One such case are emergencies where a person's life is in danger and there is either no time or no ability of the patient, or their family, to consent. As the "emergency exception" states: "if immediate treatment is required to prevent death or other serious harm to a patient, that treatment may be provided

without informed consent."(Moskop, 1999, p. 327).

Therefore, there are some justifiable cases where physicians are allowed to impinge on someone's personal autonomy, even their bodily integrity, in order to reduce or prevent greater harm to the person.

Conversely, I cannot seem to think of a case where one is allowed to infringe on someone's autonomy or bodily integrity to incur a benefit to a person, however large or considerable that benefit may be. In medicine, the procedures that are generally deemed as those that provide some benefit, for example, cosmetic surgeries, abortions, contraception, the rules for consent are as rigid as to the other procedures. Hopefully, when it comes to non-consensual cosmetic surgeries these are scenarios reserved only for horror films and not a viable option worth discussing. From this, we see that harm has a certain weight, a considerable primacy over benefit, as harm can in specific instances trespass on the personal autonomy, which is in today's society, and in medicine, one of the highest valued principles. Benefit does not procure the same privilege.

Thus, returning to Hesslow's and Bortolotti's point, it seems that the focus of medical practice is on reducing harm rather than bestowing benefit. While these are often the same, where alleviating harm is simultaneously incurring a benefit and vice versa, in cases where this is not the case harm, rather than benefit, is given primacy in medicine. I consider this as a reason enough to focus on harm rather than on benefit in my analysis.

### 4.5 Cold feet about eliminating mental disorder?

Whether the notion of harm can stand at least partly in the place of the notion of mental disorder, and how productive we can expect it to be as a partial replacement is unclear. After all, the concept of disorder seems to be entrenched in our social practices and ways of thinking about health and medicine. Therefore, eliminating the notion of mental disorder might seem controversial.

Worrall and Worrall (2001) give illustrative examples of how the question of whether something is a mental disorder seeps into the public discourse. One such example is from the 1970s when a medical practitioner in the UK prescribed smoking cessation gums Nicorette under the diagnosis of "Tobacco Use Disorder", charging the National Health Service for it. This started a notion-wide polemic, in which Norman Fowler, the Minister of Health at the time, chimed in by saying that tobacco use was just a bad habit, not a disease, and that NHS should not pay for it.

Another interesting example that they provide concerns the disorder status of stammering in the 1980s in the UK. Some claimed that stammering was correctable through education or training, and does not warrant the disease status, while others argued that it is a deviation from a norm that includes a dysfunction which would give it a disorder status (Worrall & Worrall, 2001).

A relatively recent example that brought on controversy was video game addiction that has been officially included in the ICD – 11 in 2018 (Jhee et al., 2019). The move to introduce video game addiction into the diagnostic manuals resulted in a backlash from gamers and the gaming industry who condemned such a move. As their reasons, they cited the lack of empirical evidence on the matter as well as the argument that video games help a lot of people rather than hinder them (*Debate Over Gaming Disorder Is Not All Fun and Games | Psychology Today*, n.d.). The issue whether gaming can be an addiction remains hot and contested to this day. However, the case of gaming addiction reveals interesting arguments in its rationale. Dr Vladimir Poznyak from World Health Organisation (WHO) gives reasons behind the inclusion of gaming in the ICD, as it:

> provides health professionals the possibility to identify this disorder, to diagnose it if it is present, and to link to the client all of the knowledge there is available about gaming disorder. (*WHO Expert Defends Gaming Disorder Listing*, n.d.)

Additionally, it allows them to “predict its cause and to identify the most appropriate prevention and treatment interventions”, as well as “the possibility to assign a code to allow monitoring the trends in treatment demands” which also helps in communication and dissemination of data between researchers (*WHO Expert Defends Gaming Disorder Listing*, n.d.). The article also reports that clinicians around the world observe the rise in specific symptomatology of the condition which is why “the WHO's decision to formalise it as a condition has been done so with the intention of adding cohesion and coherence to how the disorder is defined and treated” (*WHO Expert Defends Gaming Disorder Listing*, n.d.).

While these cases attest to the utility of adding gaming addiction to the diagnostic manual, we could use the same rationale behind the notion of mental disorder. The concept aids communication and research in various areas of human endeavor. It formalizes the practice and makes the treatment, as well as the theoretical background uniform and standardized. We can also see that in the case of gaming addiction the passage above hinted at a bottom-up approach where the researchers and clinicians have observed a rise in the particular symptomatology and its presence in the system. This created a need to classify and formalize the phenomena as well as the social and institutional practices around them. We could hypothesize a similar need to define and specify the notion of mental disorder. There seem to be phenomena that are of medical interest, which are handled by the domain of medicine. While there is a variety of symptomatology that these phenomena present in, they seem to share enough features to be considered under the special medical domain of psychiatry. The discourse of mental disorders and of particular diagnoses seem to unify and formalize these phenomena in order to observe, research and treat them.

Smoking, stammering and gaming are nowhere near isolated cases as similar debates about the disorder status took place concerning various other conditions,

such as alcoholism, homosexuality, PMS (pre-menstrual tension) or CFS (chronic fatigue syndrome) (Worrall & Worrall, 2001). This can be interpreted as attesting to the importance of the notion of disorder together with its implications in various domains inside and outside medicine.

Thus, even though eliminating the notion of disorder is a viable option in this meta-debate, it is one of the more controversial stances. One that I think is unlikely to come to fruition either in theory or in medical practice. Regardless, the important thing for this work is the idea that harm can withstand the attack on the notion of mental disorder and remains an element of psychiatric interest, come what may to the notion of disorder. In the next chapter, let us consider options of keeping the notion of mental disorder and how harm relates to them.

## 4.6 Conclusion

This chapter is dedicated to eliminativism, one of the prominent positions in the meta-debate on mental disorders. It is a position that advocates for the removal of the notion of mental disorder from psychiatry. I outlined the eliminativist arguments of Hesslow, Bortolotti and Ereshefsky. In focussing on harm with respect to eliminativist views I have argued that harm, even in the case of eliminating the notion of mental disorder, is an important and prominent element in psychiatry. In this context I employ the notion of 'criteria of psychiatric interest' to adapt to this meta-debate where the talk of criteria of mental disorder seems to be too narrow. I end the chapter with some remarks on the significance on the notion of mental disorder and urge for reconsideration of the eliminativist position.

# 5. The Meta-debate on Mental Disorders

## 5.1 Introduction

In the previous chapters I have outlined the problems with the traditional debate, namely the method of conceptual analysis, differing goals and different extensions of the notion of mental disorder. Then, I have presented one of the possible responses to the problems – eliminativism, which advocates for eliminating the notion of mental disorder altogether. Eliminativists consider the debate on the definition of mental disorder to be a futile endeavor which we can do without. I have shown how even if we adopt eliminativism, which does remain a controversial and seldom endorsed option in literature, harm still holds a rightful place in psychiatry.

In the following chapter I address the meta-debate on mental disorders. These are positions that are not adamant in removing the notion of mental disorder but rather are focused on reconfiguration of the debate and offering novel insights that might improve the state of the debate and bring us closer to having more fruitful discussions on mental disorders. I take the Dominic Murphy's taxonomy of positions to illustrate this move from the traditional debate to the one enriched by metaphilosophical ideas and solutions. I offer an addition to Murphy's taxonomy of positions which I call compassionatism. Compassionatism is a view that concepts such as well-being, harm, welfare, and suffering are and should be important in psychiatry. Their role is to capture the patient's perspective which I have argued earlier is one of the three pillars of psychiatric practice (see chapter 1). I defend the view of revisionist compassionatism and outline principles that such concepts should encapsulate.

## 5.2 The common denominators of the debate

Before I start with the positions of the meta-debate there is an important point that I want to lay as a groundwork for the taxonomy of the meta-debate. Dominic Murphy (2005) starts off his paper by proclaiming a very important issue:

> Everyone agrees that normative judgements play a role in assessing who is mentally ill. The difference between theorists concern whether it is our normative judgments alone that determine who is healthy or whether normative judgments must be conjoined with empirical judgments about subjects' psychology. (Murphy, 2005, p. 116)

I think the importance of this claim is largely understated and that this point is vastly unappreciated in the traditional debate. It points to the fact that the main disagreement in the debate is less pronounced and far less ominous than the debate might have us believe. Even more so, it alludes to common denominators between the contended positions in the debate, traditionally thought of as mutually exclusive. Instead of concentrating on pressure points of the debate, we can (and I believe, should) direct our attention to the points of agreement between positions to reach a common ground. This is important because it captures and makes salient where the starkest disagreement lies. Additionally, it serves as a foundation for further investigation. Murphy elaborates on this at length in his book where he defends a two-stage picture of mental disorder which can be summarized as the following. The two-stage picture

> distinguishes sharply between (i) working out when organs in the body work improperly, a scientific enterprise that psychiatry shares with general medicine, and (ii) assessing how these findings bear on our

> evaluation of lives that are affected by breakdowns. The first project is what determines that someone has a frontal lobe lesion, a depressive cognition, a genetic susceptibility to anxiety or a serotonin imbalance. The second project asks if human beings can flourish if they have such physical or psychological abnormalities.(Murphy, 2006, p. 9)

Murphy considers these two stages as separate and independent of each other. Positions which adhere to this independence and that contain an objective, scientific, value-free element Murphy considers to be objectivist. They are objectivist as long as they argue for an objectivist element as a necessary requirement and as long as this element is separate from value-laden considerations. As an example, we can take Wakefield and his harmful dysfunction account. According to Murphy's rationale, even though his view contains harm, it is not normative in the strictest sense of the term as it does not claim necessary value-laddennes of such concepts. Thus, the only real and natural adversary to Murphy's condition are pure normativists, which he mostly refers to as constructivists. As he suggests:

> Opposition to the two-stage picture comes from those who deny that we can enforce the difference between scientific assessments of humanity and the ethical, social, or political assessments (…) Opponents of the two-stage picture often believe any view about normality and abnormality in humans is irredeemably normative. (Murphy, 2006, p. 19)

To fully appreciate Murphy's claim of the two-stage picture of mental disorder, let us inspect how various position across the debate subscribe to similar elements. One such author coming from a naturalist camp is Christopher Boorse whose account of mental disorder I have covered in the first chapter. To recall, Boorse

is an avid advocate of a naturalist accounts meaning he subscribes to the view that mental disorder is a notion that can be completely realized in value-free terms in the form of biological dysfunction.

Boorse, however, differentiates between two notions: disease and illness. According to him, disease is a scientifically objective notion which he defines as a function that is statistically below the norm concerning the age of an organism (Boorse, 1977). However, disease is an illness only:

> if it is serious enough to be incapacitating, and therefore is (i) undesirable for its bearer; (ii) a title to special treatment; and (iii) a valid excuse for normally criticizable behavior. (Boorse, 1975, p. 61)

We see here that Boorse proposes two alternative notions, which are in most part and in general, synonymous with the notion of mental disorder.[16] Disease is a value-free notion while illness is evaluative. According to Boorse, their roles in the medical practice differ as well since disease is a theoretical, biological notion that "applies indifferently to organisms of all species" (Boorse, 1975, p. 56). Conversely, illness is a subcategory of diseases "that have certain normative features reflected in the institutions of medical practice" (Boorse, 1975, p. 56). Accordingly, not all diseases are illnesses, but only serious ones, that are undesirable by the person, that deserve to have special treatment and those that validly excuse normally criticized behavior, as I have mentioned above. These all seem to be criteria belonging to the normative side of the debate. What is important to notice is that Boorse recognizes the normative roles the term illness plays but rather than conflating it with disease he emphasizes the distinction.

[16] Many authors conflate the notions of disease, illness, disorder and take them to mean roughly the same thing.

Now other authors, particularly of the normative camp would argue that the distinction is maybe forced, and that one notion suffices, it being disease, which is inherently value laden. Here Boorse, who is a hard-core naturalist pertains to the idea that values enter at least at some point with regards to the concept of mental disorder in terms of illness.

Thus, there is an objective, scientific, descriptive element which purports to be value-free as much as possible, and then there is a normative element that is value-laden and freely contains normative components. In the context of psychiatry, and in the context of diagnostic manuals that we are discussing, all of which are institutionally entrenched, the relevant elements that we can assume are at play at determining whether someone is disordered are two; one that is descriptive, scientific and one normative, evaluative.

These two elements can be found in a prominent hybrid account I have outlined earlier, Wakefield's harmful dysfunction (Wakefield, 1992, 2014) that contains the notion of harm as an evaluative element and the notion of biological dysfunction as an objective, scientific and value-free element (see chapter 1).

Outside of the debate on mental disorders, the descriptive and evaluative component are also found in authoritative diagnostic manuals such as DSM and ICD. Here is the definition of mental disorder in DSM:

> A mental disorder is a syndrome characterized by clinically significant disturbance in an individual's cognition, emotion regulation, or behavior that reflects a dysfunction in the psychological, biological, or developmental processes underlying mental functioning. Mental disorders are usually associated with significant distress or disability in social, occupational, or other important activities. An expectable or culturally approved response to a common stressor or loss, such as the death of a loved one, is not a mental disorder. Socially deviant behavior

> (e.g., political, religious, or sexual) and conflicts that are primarily between the individual and society are not mental disorders unless the deviance or conflict results from a dysfunction in the individual, as described above. (American Psychiatric Association, 2013, p. 20)

Here we see an element of biological dysfunction which assumes the role of scientific objectivity, and the notion of distress or disability (the harm criterion) which is an evaluative, value-laden component.

The controversy so far was with whether these elements were necessary and sufficient conditions for the definition of mental disorder (Schwartz, 2014). However, in the novel turn that the debate has taken the focus moves towards novel ways of framing the dichotomy between value-free and value-laden criteria. Ereshefsky (2009, p. 221), instead of engaging the debate on the necessary and sufficient conditions of disorder shows how the efforts of determining necessary and sufficient conditions are misguided and makes as his focus "considerations that are central in medical discussions". As I have already mentioned when discussing eliminativism, Ereshefsky notices that throughout the discussion on disorder, two considerations were especially salient: state descriptions and normative claims. State descriptions are defined as "descriptions of physiological or psychological states"(Ereshefsky, 2009, p. 225), which avoid normative elements, like natural, normal or even function, as much as it is possible. Presumably, this element purports to be as value-free as possible and is a placeholder for an objective, scientific element. Normative claims, on the other hand, "are explicit value judgments concerning whether we value or disvalue a physiological or psychological state" (Ereshefsky, 2009, p. 225) and represent normative, value-laden considerations or elements. Here we again see the two-partite divide once again.

Now, based on these we can draw a floorplan on which to base our inquiry. Inspired by Murphy's two-stage picture, supported by the positions across the

discussion, and Ereshefsky's eliminativism, or as I would call it cautious agnosticism about the concept of mental disorder, I propose the following.

There are elements that are present in the debate that are recurring across positions which are rightfully deserving of our investigation. In broad strokes, I argue that there are two elements of psychiatric interest, a) normative, value-laden element, and a b) scientific, objective component.

Before I begin, I should put a caveat in place. I am not arguing here that the definition of mental disorder should contain a normative and a descriptive element, as that is simply what a hybrid position pertains to in the original debate between normative and naturalist accounts, or might even fall under kinds of objectivism that Murphy argues for. Instead, I am arguing that a normative and a descriptive component, as it seems, will be of great importance in determining whether some condition is of medical, and specifically psychiatric, interest. Returning to the idea I have hinted at in the previous chapter, instead of criteria of mental disorder which are not as applicable to this meta-discussion, I propose the *criteria of psychiatric interest*. These are the criteria that do not have to strictly fall inside the notion of mental disorder but are criteria particularly salient, important, and reasonably worthy of our analysis in psychiatry. These are the criteria that play key roles in assessment, diagnosis, treatment, research, and alleviation of psychiatric conditions.

One worry that might be raised against such the formulation of 'criteria of psychiatric interest' is that while it is somewhat broad, it might seem all too vague. After all, there might be, and probably are numerous criteria of interest for psychiatry at various levels and domains of psychiatric practice which makes the job of teasing apart or narrowing down those important that much difficult.

This could pose a legitimate concern. However, what I aim to argue, and what I have hopefully shown so far in this section is that claims about these two criteria and their interaction is reducible to two. Thus, the multitude of criteria that arise at

various point in analysis, both inside or outside the scope of the definition of mental disorder can fall into either of these two camps: descriptive or normative elements of psychiatric interest.

Now, I remain agnostic on whether these elements are found in the definition of mental disorder, outside of the definition or enacted on in some other place in the process of diagnosis, as well as in the scientific context and practices surrounding it. All I am willing to argue at this point is that both normative and descriptive consideration will remain to be crucial, and essential in determining whether a person suffers from a psychiatric condition and the steps that will be taken to alleviate it. In other words, there are two groups of elements which are of great psychiatric and medical interest: one normative, other descriptive.

I consider this to be in large part compatible with Murphy's account. Even more so, while Murphy argues for a two-stage picture in mental disorder, I bring forward a weaker claim. Instead of the narrow scope of applicability just to the concept of mental disorder, I apply the two-partite view to the psychiatry in terms of elements of psychiatric interest. This is compatible with eliminativism, and with an array of positions. This is favorable for my thesis because my goal is not to adjudicate between normativism and naturalism in the traditional debate. My aim is to explore harm - how harm factors in psychiatry and how it is configured. Bearing this in mind, let us turn to Murphy's taxonomy of position for further analysis.

### 5.3 Murphy's taxonomy of positions

As I have disclosed in the previous chapter, the traditional debate on mental disorder faces serious problems. As a response, novel approaches have arisen to remedy the underlying issues of the traditional debate. Even though each of the approaches aims to improve the prospects of the traditional debate the variety of the positions and differing theoretical set-ups might lead to incommensurability and unintelligibility of positions in relation to each other. This has the potential to end in a multitude of

methodological and conceptual novelties that might be even further from what was initially aimed regarding the notion of mental disorder. Instead of improvement it might bring confusion. Different from the traditional discussion which has positions and rationale, methodology laid out for it, the meta-debate seems to lack and be in the need of a bit of systematization. Thus, grounding the meta-debate and systematizing all the nuances of the positions seems to be a worthwhile endeavor. One such effort has been undertaken by Dominic Murphy. The specifics of his taxonomy is that he builds on, and includes the main positions of the traditional debate as foundational on which he applies peculiarities from the meta-debate. Thus, in the following I lay out Murphy's taxonomy of the debate. I outline how the positions covered so far can be integrated into this system. However, Murphy in his book does not provide his systematization with the elaboration I think it deserves which I aim to remedy in the remainder of this chapter.

To begin, Murphy (2005, p. 117) cites Philip Kitcher (1997, p. 209-209), whose dichotomy between objectivism and constructivism is considered to be foundational in medicine:

> Some scholars*, objectivists about disease*, think that there are facts about the human body on which the notion of disease is founded, and those with a clear grasp of those facts would have no trouble drawing lines, even in the challenging cases. Their opponents, *constructivists about disease*, maintain that this is an illusion, that the disputed cases reveal how the values of different social groups conflict, rather than exposing any ignorance of facts, and that agreement is sometimes even produced because of universal acceptance of a system of values.

This dichotomy between objectivists and constructivists is evocative, and

analogous to the dichotomy of naturalism and normativism found in the traditional debate on mental disorders. However, there are some differences that should be highlighted. Murphy lumps into objectivists all accounts that argue for an objectivist or naturalist element as separate and independent of constructivist or normativist. Thus, hybrid positions, such as Wakefield's fall into his objectivist camp as he argues for a notion of dysfunction that can be separated from the normative element of harm. Under constructivism Murphy places all of the theories that argue that determining the disorder status rests essentially on normative judgments alone.

He postulates that there can be two modes, or variants, of each of the positions, either constructivist or objectivist. These are *conservative* and *revisionist*. A conservative view "says that our concept of illness is correct and should constrain a theoretical picture of health and disease worked out by scientists and clinicians." (Murphy, 2005, p. 117). In other words, in asking ourselves what a certain concept amounts to, we are doing conceptual analysis of the notion (Murphy, 2005, p. 117) – searching for necessary and sufficient conditions of its application. This approach can be viewed as descriptive in a sense that it pertains to determining how the concept *is* used.

Conversely, revisionists think that while our concepts, and our intuitions about the concepts of health and disease are necessary and at the forefront in the beginning of our exploration, "they should not constrain the inquiry" (Murphy, 2005, p. 117). In looking for what particular concept entails we might find that "our concepts of health and disease are inadequate when judged against the facts, or that they fail to meet the requirements of science in some way." (Murphy, 2005, p. 117). If and when it comes to that it is our prerogative to tailor, the concept however we deem it best to fit with the requirements and desiderata that a certain concept commands. This can be considered as a normative approach since instead of exploring how the concept is used, we are pondering how the concept *should be* used.

Thus, conservativism and revisionism can be seen as two kinds of modalities each of the positions, either objectivist or constructivist, can take. Let us now briefly go through all the permutations of the positions to capture in its fullness the nuances of the positions that Murphy brings forward.

### *5.3.1 Conservativist objectivism*

First position to consider is *conservativist objectivism*. Murphy argues that objectivists are often conservative in that "they believe that common sense incarnates what philosophers call "folk psychology", which is a set of everyday unscientific theories about the mind." (2005, p. 117). Conservativists think that the folk psychology determines what is mental disorder and that scientists and clinicians should adhere to it and look for things in the world that correspond to that image of mental disorder. Murphy considers this to also be one of the serious downsides of conservative objectivism as he thinks that "this understanding of science's relation to common sense should be rejected."(Murphy, 2006, p. 20). His book *Psychiatry in the scientific image* (2006) is relentlessly devoted to arguing for this breakaway of psychiatry from common sense ideas about psychiatric concepts, i.e. folk psychology.

One such author he discusses, and criticizes, is Wakefield. Recall, Wakefield defends a harmful dysfunction account meaning that "our concepts of both mental and physical disorders involve two individually necessary and jointly sufficient components" (Murphy, 2006, p. 37). Wakefield is traditionally a hybrid theorists in that mental disorders contain normative and naturalist element.

In Murphy's taxonomy Wakefield is considered an objectivist because he defends an objectivist component which is independent from the normative one, as I have indicated earlier. Wakefield considers folk psychology to be authoritative with regards to the intuitions about what mental disorders are. As Murphy notices: "Wakefield is *criticizing* the scientific, theoretical picture of mental disorder by an appeal to intuitions. This theme comes up time and time again in Wakefield's

writings." (Murphy, 2006, p. 52). Wakefield thinks that "our current ignorance of human psychology is beside the point, because "we do not have to know the details of evolution or internal mechanisms" to make judgements that some inner mechanism is malfunctioning" ((Wakefield, 1997a, p. 256) as cited in (Murphy, 2006, p. 44). Instead, Wakefield seems to think our intuitions capture whether someone is functional or dysfunction based on the "surface features" that reveal that (dys)functionality ((Wakefield, 1997a, p. 256) as cited in (Murphy, 2006, p. 44)). This means that scientific investigation about mental disorders come second to our (folk) intuitions about them. While this might have been a starting point of psychiatry in the past, when scientific investigations were few and far in between, in contemporary medicine this would largely be considered a controversial way of thinking about disorders. Imagine if in discussing and diagnosing cancer we were relying on what our intuitions tell us about cancer rather than looking at scientific evidence related to it.

Wakefield is not alone in this, as other objectivists, most notably Boorse, also "adduce everyday linguistic usage and commonsense intuitions as evidence"(Murphy, 2006, p. 52). While Murphy comes up with more elaborate attacks to traditional objectivist accounts, this is beside the current point. The gist of conservativist objectivism is giving priority to the folk intuitions and moulding scientific ones to their liking. Murphy is ardently against science playing such a perfunctory role in assessing mental disorders. While the elements of dysfunction are viewed as embodying scientific objectivity, the way they are approached in the objectivist part of the debate does not reflect that ethos. Thus, Murphy rejects conservativist objectivist view and advocates for revisionist objectivism.

### *5.3.2 Revisionist objectivists*

Revisionist objectivists say that "facts about physiological and psychological functioning obtain regardless of how we think about disease"(Murphy, 2005, p. 117). This means that our intuitions and common sense can get things wrong. Not only that

but they should not be decisive factors in how we consider and approach our scientific concepts. According to revisionist objectivists, disorder and health are "features of the world, discoverable by biomedical investigation."(Murphy, 2005, p. 118). Thus, medical discovery and clinical conceptions take precedence over folk psychological concepts and are not curtailed by them. Murphy declares himself a revisionist objectivist defending his position "based on the belief that the objectivist program only can be carried through if psychiatry and cognitive neuroscience are synthesized" which "violates some commonsense views about mental illness" (Murphy, 2006, p. 21) which is the revisionist part of his account.

Another author that can be considered revisionist objectivist is Peter Schwartz. He rejects conceptual analysis by arguing that the idea of concepts consisting of necessary and sufficient conditions is a bit outdated and not particularly useful. He is of the opinion that "accounts should be seen as proposed *new* definitions, or groups of new definitions that can do much of the same work that was done by the vaguely defined concept of disease in the past."(Schwartz, 2007b, p. 60) but instead of discovering them, we decide on them. Thus, Schwartz opts for a prototype theory view of concepts where cases are sorted based on their resemblance or divergence from some prototype, as he argues that that is in fact how our conceptual apparatuses work based on psychological studies (Schwartz, 2007b, 2014). Additionally, he contends that "Using prototypes and similarity metrics has the advantage of being more in keeping with the actual use of 'disease' and other concepts." (Schwartz, 2007b, p. 60). According to him scientists are more partial to the Wittgenstein's model of "family resemblance" which corresponds to prototype theory of concepts where things fall under concepts within degrees of similarity, rather than being all-or-nothing affairs (Schwartz, 2007a, p. 57). For example, thinking about the concept of birds, prototypical examples are birds like pigeons, swallows,

ravens, while less prototypical and further from these are penguins, chickens and so on.

Bearing all this in mind, Schwartz offers an alternative to conceptual analysis in the form of philosophical explication of the notion of mental disorder (Schwartz, 2007b, 2007a, 2014). Explication is defined as a "process whereby a vague, informal concept, either from everyday life or from more regimented contexts such as scientific contexts (but in earlier stages of development), is given a more exact, often formalized formulation." (Dutilh Novaes, 2018, sec. 1.1. Carnapian explication). It is a project inspired by Carnap and Quine, pioneers of the philosophical explication (Carnap, 1971; Quine, 2013). Here is how Quine describes his motivation of explication:

> We do not claim synonymy. We do not claim to make clear and explicit what the users of the unclear expression had unconsciously in mind all along. We do not expose hidden meanings, as the words 'analysis' and 'explication' would suggest; we supply lacks. We fix on the particular functions of the unclear expression that make it worth troubling about, and then devise a substitute, clear and couched in terms to our liking, that fills those functions. (Quine, 2013, p. 238)

Thus, philosophical explication has a strong pragmatic undertone and is integrated in scientific and other practices relevant for the notion. This means that analysis of the term to capture its essence is not the goal. Rather, the goal is to devise a concept applicable for a particular situation or scientific context in which it is employed.

Carnap illustrates the process of explication through a famous example of the notion of 'fish' which has been in scientific contexts replaced with the concept of 'piscis' (Carnap, 1971). 'Fish' refers to all kinds of sea creatures that live in the water

which vaguely resemble each other. In folk and everyday discourse, this might be practical enough to be utilized without considerable problems. However, on this understanding of the notion of 'fish', whales, dolphins, stingrays, sharks and basses all fall under the same concept, which in zoology or other scientific disciplines is too wide and imprecise of a category to be utilized pragmatically. Therefore, a need arose to specify the concept of 'fish' by introducing the concept of 'piscis' which would exclude whales and seals and similar sea creatures, and would be unambiguous, more adequate, and more precise than 'fish' (Carnap, 1971). In this scientific context, the notion of 'piscis' would thus replace the vague concept of 'fish' (Carnap, 1971).

Returning to the notion of mental disorder, the aim of the explication would be to specify the notion and to make it more precise and explicit. As Schwartz recognizes, one of the motives of trying to define mental disorders in terms of necessary and sufficient conditions is to determine whether the borderline cases fall under the notion or outside the notion of disorder (Schwartz, 2007a, p. 60). Opting for philosophical explication turns this problem into a decision rather than a discovery (Schwartz, 2007a, p. 60) meaning we can pragmatically tailor the notion of mental disorder according to best reasons and arguments.

This however does not mean that the traditional debate on the notion of mental disorder has been futile as Schwartz recognizes the important role it plays in the process of explication:

> From this perspective, the long debate between normativism and naturalism – including the discussion of the four theories described above – serves a crucial role by displaying the advantages and disadvantages of the various approaches. Choosing an account becomes not so much a hunch about which theory is correct but instead a choice of which theory to clarify and apply. (Schwartz, 2007a, p. 60)

Therefore, the current debate between naturalism and normativism provided us with an abundance of theoretical material in terms of different accounts of mental disorder, their peculiarities, and conceptualizations as well as the possibilities of practical utilization. This provides an opportunity to choose among theories and to specify them further to fit particular purposes. This can also result in pluralism of concepts that can be readily employed in various circumstances. As Schwartz concludes:

> Once we accept that the project of defining 'disease' is a constructive one, and that definitions do not have to be stated according to the classical view, there is room for many approaches. Where there was once stalemate, let there be variability and free choice.

As such, the approach is more adaptable and applicable in various circumstances and opens possibilities for various new and innovative approaches to the concept of mental disorder.

Furthermore, the approach of philosophical explication undoubtedly brings with it some changes both in methodology and in the resulting concept of mental disorder. Schwartz points out several novelties and consequences of adopting the philosophical explication. One such implication are deviations from current usage of the concept (Schwartz, 2014, p. 578). While in conceptual analysis, the idea is to describe and discover how various users of the language use the concept, and how it fits their intuitions, in philosophical explication deviations from standard use are expected since "The currently used concept is not clearly defined or it is defined in ways that are problematic, and thus a new definition is desirable." (Schwartz, 2014, p. 578). The whole aim of explication is to specify the concept and to make it clearer which will understandably result in some differences in the use. This captures the revisionist aim of his account.

Another consequence of employing philosophical explications is that it will impact the traditional discussion on mental disorders, as according to Schwartz (2014) it will reframe it. Since the discussion is about which concept is the true one, there can be only one that captures all the intuitions. The competing sides of the debate, naturalists and normativists, aim to uncover that concept. In philosophical explication, however, it is not the matter of which concept is the true(st) but which best serves a particular purpose (Schwartz, 2014, p. 579). Thus, the debate takes a pragmatic turn to refocus on which concept to adopt (Schwartz, 2014, p. 579).

Finally, and somewhat connected to the previous point, since the debate shifts to normative instead of a descriptive project additional burdens are in "identifying and justifying the specific theoretical roles these terms will play" (Schwartz, 2014, p. 579). In descriptive projects, in projects that look for the one, true account of mental disorder, the implicit idea is that one argues for the true one as that is the goal everyone in the discussion seems to be aiming at. This concept would encompass all the relevant intuitions and desiderata of disorder. However, in philosophical explication the goal is to find the most appropriate, most useful in a specific context and thus needs further justification why a particular rendition of a concept should be used over another.

We can also interpret this as a downside of the project of philosophical explication because the notion of mental disorder loses a 'truth-maker' in the world but becomes a matter of convention. It ceases to be 'discovered' but is devised. While this might ease the burden of defining the concept in getting to the 'true' or 'right' one, it makes it especially vulnerable to relativism. The consequences of the definition of mental disorder, as I have pointed out in the first chapter, are manifold and serious. The project of explication must insure itself against "anything goes" attitudes, which are sometimes attributed to relativism.

Furthermore, Cooper (2020) points out two additional drawbacks of

revisionist projects; one being that as the contexts and environments change, it will be hard to predict the future usefulness of the concept. As the concept is devised with certain aims and purposes in mind once these change or become obsolete the concept itself might be brought into question.

Another downside is that "when terms get defined differently for different projects, this can lead to coordination problem where such projects meet" (R. V. Cooper, 2020, p. 151). For example, mental disorder is used in various contexts such as law, public policy, medical insurance, and other contexts. Constructing a specific notion of mental disorder just for the purposes of psychiatry has the potential to cause problems and confusion in these other contexts that might work with a different notion of disorder. It can result in a multitude of meanings in an array of disciplines which would not be able to communicate to each other, which is something we might want to aspire to avoid.

As I have mentioned in the first chapter, the notion of mental disorder is employed in many different areas and having a different, specific definition for each of the contexts, as well as further breaking into narrower concepts might create more problems and inabilities between disciplines to communicate.

Schwartz (2007a, p. 60) offers his own taxonomy of the positions of the traditional debate. He differentiates between value-requiring, non-value requiring, dysfunction-requiring and non-dysfunction-requiring theories (Schwartz, 2007b, p. 49). As the names suggest, the theories in the discussion are divided between these positions depending on the content of their theories. For example, Wakefield's harmful dysfunction account is both value-requiring and dysfunction-requiring account, Boorse's is dysfunction-requiring, non-value-requiring, and so on (Schwartz, 2007b, p. 49). Schwartz (2007b, p. 60) defends dysfunction-requiring theories as they do "a good job making sense of prototypical cases of disease and distinguishing them from healthy, undesirable conditions.". He goes through a couple of prototypical

cases where he suggests that dysfunction-requiring accounts do a better job than others. He concludes with what can be considered the gist of his project:

> Once we accept that the project of defining 'disease' is a constructive one, and that definitions do not have to be stated according to the classical view, there is room for many approaches. Where there was once stalemate, let there be variability and free choice (Schwartz, 2007b, p. 61)

Thus, Schwartz offers a methodological novelty in comparison to the traditional discussion while siding with objectivists about disorder and arguing for explication of such concepts.

### *5.3.3 Conservative constructivism*

Constructivists, similar to normativists (see 1st chapter), can agree to some biological facts about disorders, but what in their opinion drives the notion of disorder, and is essential to it, are values, i.e. social norms (Murphy, 2005, p. 118). Their rationale is that we live in a society that adheres to some social norms. Those who transgress those norms are either considered to be immoral or ill (Murphy, 2005, p. 118). The biological or psychological facts come second to moral norms, and they are used to justify the previously determined decision based on transgression of social norms (Murphy, 2005). The main differentia specifica between normativism and naturalism (i.e. constructivism and objectivism) is that the former consider disorder essentially, necessarily value-laden. As Murphy explains:

> A constructivist can concede that we look for distinguishing features in the biology or psychology of the deviants. But a constructivist will say that we do this only because we first decide on other grounds that these people

> are mentally ill and that we then cast about for something about them that we can medicalize. (2006, p. 24)

In this view, scientific facts are not independent but are heavily influenced by our judgments about conditions. It gives the impression of inseparability of facts and values and evokes hard skepticism of its disentanglement.

When it comes to constructivists, there are two variants as well. First are *conservative constructivists*, who "aim to uncover the value-laden bases of our folk psychopathology" and who say "there's no more to psychopathology than behavioral patterns that violate social norms" (Murphy, 2005, p. 118). Their aim would be to make explicit our social norms and the way psychopathology confers to them. This would be a descriptive project of describing normative terms in the discussion. It is based on the belief that our ideas on mental disorders are normative and inseparable from scientific facts as well as that these norms just track social currents on what is viewed as pathological. It is exploring how things are meaning it is a descriptive project. One such example is Rachel Cooper's belt-and-brace approach. As I have mentioned earlier, in criticism of traditional debate (see 2nd chapter) Cooper asserts that the concept of mental disorder is a human kind term, borrowing from Hacking (1996). This means that it is changing due to social and political changes at a rate that makes conceptual analysis of these concepts somewhat futile. Because of this, the necessary and sufficient conditions are hard to capture and even if they were captured, they would not stay as such for long. Mental disorder, writes Cooper (2020) is one such term.

Her critique of the traditional debate is based on this insight of the changeability of concepts and efficacy of investigating them through conceptual analysis. Cooper's solution to the problem of elusiveness of the concept is to belt-and-brace conceptual claims. Her idea is to fortify the existing and consistent ideas in the

debate, which would withstand the test of conceptual shifts. As Cooper explains, when one is not sure if their belt is going to be able to hold their trousers on their own, and their braces are also unreliable, then the best course of action is to belt and brace one's pants to maximize the chances of them not falling off. Similarly, we should identify and strengthen the conceptual claims concerning the notion of mental disorder. As such, anchoring some fairly consistent and robust claims would provide resistance to the conceptual ebbs and flows of social and political currents.

With this in mind, Cooper argues that "normative judgements are involved in determining whether a condition is a disorder", and that "no account of disorder can be given that appeals only to biological facts, and that normative judgements will be necessary." (R. V. Cooper, 2020, p. 153). This is the claim she considers robust, which has the potential to persist through conceptual shifts. In her terms, this is the claim that she "belt-and-braces".

### *5.3.4 The flavors of normativity*

As I have already mentioned in the second chapter, there are different levels of normativity when it comes to the concept of mental disorder (Kingma, 2014). This is also the idea on which Cooper (2020) builds her claims. She outlines two ways in which values enter the considerations of mental disorders. To one of these she refers to as the threshold problem. Cooper notices that our thresholds for when some condition becomes a disorder vary as the "boundaries of some disorders are currently set at the extreme tails of normal distributions, while the boundaries of others are much closer to the mean."(R. V. Cooper, 2020, p. 154). According to Cooper, these are not grounded on some objective and statistical factors but on the considerations such as cost-benefit analysis of alleviating the condition where "function of the ease and expense of treatment" play important roles (R. V. Cooper, 2020, p. 154). For example, as treatability of conditions such as hypertension and ADHD improved and became relatively accessible, so the threshold moved to include more instances of the

condition. As such, it seems that where we draw the threshold is at least in part determined by our values and goals, which is a hallmark of normativity.

Another way in which values enter into our considerations is what she refers to as the 'location' problem (R. V. Cooper, 2020, p. 155). In analyzing some conditions of medical interest, we can locate the issue inside or outside of the individual. For example, in people with pollen allergies we can treat the allergies, or the person can move to another place or avoid locations where there are large amounts of allergens. As Cooper notices, "we tend to think of the location of the problem differently depending on whether we think the individual or the environment should be altered." (2020, p. 156). For example, in allergies we locate the problem as internal to the individual and alter or treat the individual, while in the examples of poison or toxic aerosols, we locate the problem in the environment, so we move the individual away from the toxic substances.

Especially instructive is the case of homosexuality, where at first, before 1970s, homosexuality was widely considered to be a mental disorder meaning that the problem was in the individuals. However, the situation changed, in large part due to gay activists, who insisted that the problem is with the societal disapproval of homosexuality, not with homosexuality itself (Bayer, 1987). The psychiatric authority had a sympathetic ear for the activists and realized that there might not be something inherently pathological in homosexuality. As a result they demedicalised homosexuality by removing it from the list of mental disorders (Bayer, 1987; Drescher, 2015).

The lesson here and the point which Cooper (2020, p. 157) brings forward is that where we locate the problem rests on various considerations, such as whether we can change the person or the environment, what kind of change is reasonable and so on, which in turn rests on a number of factors, including social and political ones.

Thus, the matter is determined by some factors outside of the descriptive,

scientifically objective facts about conditions, as we have seen the problem location shift in the case of homosexuality, from being an internal one to being a problem in the environment, i.e., society.

These two instances illustrate Cooper's point that normativity can be found at various levels of the concept of mental disorder, which makes normativity ever-present in the notion of disorder. This idea is also the one that Cooper thinks is robust, and that normativity will for some time be present when it comes to the notion of mental disorder, albeit at different levels of the concept. This can be considered as a hallmark of conservative constructivism as Cooper's project is descriptive meaning she maps out conceptual changes and issues without taking a normative stance, and it argues that the concepts are and will be essentially normative, i.e. constructivist. Her belt-and-brace is not so much a vindication of the traditional view of normativity that was defended in the original debate but the crux of her argumentation is that on traditionally considered naturalist or objectivists side, normativity is essential and unavoidable.

While I agree with much of what the Cooper is saying, I think that most of the positions in the traditional debate would actually agree too, both on the normativist and on the naturalist side of the debate. In their recent work, Cristina Amoretti and Elisabetta Lalumera (2022) suggest this by discussing what the belt-and-brace approach does in terms of the traditional debate on mental disorders. They show that the normativity at different levels, such as the one that Cooper proposes, can be viewed as compatible with an array of naturalist accounts. As such, it does not seem to do much in vindicating normativity, or, it could be argued, it is the normativity that is familiar and accepted by naturalists. It would mean that the concept of mental disorder is *no more* normative than other scientific concepts, which I believe is a claim that many naturalists can comfortably live with.

To illustrate, Amoretti and Lalumera (2022) introduce the dichotomy between

*weak* and *strong* normativism. According to them, "the disease concept is strongly normative if an evaluative concept, such as disability, action failure, harm, suffering, unluckiness, or undesirability is one of its explicit components, or a necessary criterion for its application." (Amoretti & Lalumera, 2022, sec. Weak and strong normativism). This roughly corresponds to what normativism in the traditional debate on mental disorder amounts to. Both normative and hybrid authors can be subsumed under such a view – Cooper, Reznek, Nordelfelt, Megone and Wakefield (see chapter 1). Even in those not adhering to the method of conceptual analysis, Amoretti and Lalumera claim, authors who consider values and goals as "a central explanatory tool of a general theory of diseases" (Amoretti & Lalumera, 2022, sec. Weak and strong normativism) would be considered as strong normativists. In short, according to them, if a normative component is explicit in the definition of a disorder, or if a normative component is necessary, or a key explanatory component of a theory of disorder, then this is a case of strong normativism. For example, Wakefield's harmful dysfunction view of mental disorder, or the view of mental disorder in DSM IV (1994) as including the elements of both dysfunction (biological, psychological or social) and harm (distress or disability) can be considered in their view as strongly normative. In contrast, a concept of disorder is weakly normative:

> if no evaluative concept explicitly figures as a component of the definition of disease, but some value-laden concepts or judgments may intervene in the operationalization of some of such components. (Amoretti & Lalumera, 2022, sec. Weak and strong normativism)

Under weak normativist positions they consider cases of what they call 'value intrusion'. This is problematized by Cooper in her discussion of threshold and location problem that I have outlined earlier, as well as somewhat earlier in Kigma (2014). Amoretti and Lalumera besides Cooper's also discuss Elsejin Kingma's (2014) paper

where she lays out different ways in which normativist and naturalist position interact and can be opposed. In short, Kingma introduces four such domains:

> (1) "health" and "disease" as ordinarily used, (2) theoretical or conceptually clean versions of "health" and "disease," (3) the operationalization of dysfunction, and (4) the justification for a given operationalization of dysfunction. (Kingma, 2014, p. 591)

She notices that the opposition between naturalism and normativism in the traditional debate corresponds to the second level "theoretical and conceptually clean versions of "health" and "disease"". This standard clash of the descriptive and normative is seen in the debate on the notion of mental disorder (see chapter 1). However, besides this kind of normativity, there are two additional levels of normativity that can be observed. These can be found in the criterion of dysfunction, usually taken to be naturalist, as well as descriptive aspiring to be scientific and objective. These third and fourth level of normativism rely on a subset of arguments that "question whether, assuming that an analysis of disease as dysfunction is correct, the concepts of function and dysfunction can themselves be defined or operationalized in value-free terms." (Kingma, 2014, p. 594).

These kinds of normativity in terms of value-intrusion of the naturalist component are generally recognized by naturalist accounts. Thus, it is not clear what Cooper's account of belt-and-bracing normativity does in terms of the debate on mental disorders. It does not adjudicate the debate, or pushes in either the normativist or naturalist direction. Granted, maybe it is not supposed to. But it also does not give us clues, or insights to proceed further in the meta-analysis of the debate since at first glance one might suppose it advocates normativism in the debate, while at a closer inspection it seems to be compatible with positions across the board.

Correspondingly, it is a bit unclear what then Cooper's idea of belt-and-bracing normativity does for the concept of mental disorder. It may outline one instance that aspires to be robust and stand the test of time. Here I recognize the merit of belt-and-bracing normativity. The awareness of normativity tends to fall into the background, especially on the naturalist side of the debate. However, it is an important lesson that normativity operates even when it comes to naturalist criteria. I believe it is important to be cognizant of the various ways and levels that normativity plays a part in both sides of the debate, naturalist especially. This idea I see as an invaluable to be at the forefront of our considerations about mental disorder and conditions of psychiatric interest in general.

Nonetheless, belt-and-bracing normativity does not seem to justify using and keeping the notion of mental disorder. While Cooper seems to advocate for keeping the notion of mental disorder, her ideas about how normativity at different levels would aid doing that seem to be a bit underdeveloped.

Cooper is thus considered a conservative constructivist since her account tracks the movements of the harm criterion and presupposes these as relevant for our definitions in psychiatry. Since the consensus on harm is 'breaking down' our conceptual systems should track this change. This reveals a bottom-up process from the folk and expert intuitions to conceptual systems that we keep in place.

### *5.3.5 Revisionist constructivism*

Besides conservativists, there are *revisionist constructivists*, who instead of having a descriptivist attitude of tracing how our society tracks disorder, assume a normative one and tend to criticize the social norms that are in place. They argue that "concepts of health and disease are used to medicalize behavior that is really just socially counter-normative." (Murphy, 2005, p. 118) and that power relations and struggle between them is foundational of our nosologies. Revisionist constructivists want to show how "we have adopted a theory of psychological malfunction that just covers

up the real power relations at work."(Murphy, 2005, p. 118). A well-known example of revisionist constructivist in literature may be Michel Foucault and his work *Madness and Civilization* (1971) where he criticizes the way psychiatry adopted values as medically authoritative. Foucault's work describes history of psychiatry as misuse of medical institutions for ideological purposes. The remnants of such practice, according to Foucault are still very much alive today in contemporary psychiatry. His work is a criticism of such practice.

Returning to Murphy's taxonomy (2005, p. 118), he goes on to say that objectivism and constructivism can be combined. This is similar to normativism and naturalism being combined in a single theory - hybrid theories (see 2st chapter). There are various ways of combining them, as Murphy (2005, p. 118) outlines: "one can be a objectivist about bodily disease, but a constructivist about psychiatry". In literature, an example of this would be Thomas Szasz' essay "The Myth of Mental Illness"(1960). According to him, there are no such thing as mental illnesses – they are a myth. What we call mental illness is either a) a brain disease that will in time be discoverable by medical science, and thus can be equated with bodily disorder, or b) a deviation from a norm "that must be stated in terms of psychological, social, ethical and legal concepts."(Szasz, 1960, p. 114). The latter cannot be remedied by medical means since "it seems logically absurd to expect that it will help solve problems whose very existence had been defined and established on nonmedical grounds."(Szasz, 1960, p. 115). Thus, Thomas Szasz tends to be an objectivist about bodily illness and constructivist about psychiatry. One could also be constructivist about particular conditions, or periods of history of psychiatry and objectivist about other (Murphy, 2005, p. 118). I will not go into this at lengths, let us return to the fourfold distinction I have outlined earlier. Murphy argues against constructivism in favor of revisionist objectivism.

## 5.4 Murphy's critique of constructivism

He poses the following as the challenge for constructivism:

> why do we not regard all deviants as mentally ill, if mental illness is just a matter of violating norms? And what makes us change our mind about some norms so that they become less important to us over time and the violators shift status from mentally ill to merely unusual, or even come to be admired? (Murphy, 2005, p. 119)

Objectivist's answer to this is that people deviate from social norms because there is something faulty with them which is signaled by the deviation (Murphy, 2005, p. 119).

Having this in mind, Murphy outlines two serious problems that constructivism faces. One such problem is that we do in fact seem to tease apart the social deviant from the pathological. However, constructivism does not have tools nor explanation on what account we are able to do that. If everything there is to mental disorder are social deviations, it becomes hard to establish and give reasons why some behavior, processes or functions are deviant, while others are pathological. Thus, Murphy (2005, p. 119) argues that constructivism faces problems in "its inability to explain everyday distinctions between the pathological and the merely disapproved of", as he continues "No constructivist has explained why we call violations of some norms but not others a disorder.". He explains that we do not think of "chronic rudeness, fascist ideology and physical ugliness" (Murphy, 2005, p. 119) as disorders, even though we socially disapprove of them.

Another problem for constructivism, connected to the previous point, is the relativist objection. Murphy (2006, p. 27) argues that constructivist "cannot distinguish moral and political disputes from psychiatric ones". We, however, claims

Murphy have the ability to do that. Murphy thinks that we can discern this and even "criticize putative diagnoses on the ground that they are politically motivated and hence fraudulent" (Murphy, 2006, p. 27). Murphy reminds us of infamous examples from psychiatric history which are characteristic of this ability for example, drapetomania and masturbation. If it were up to constructivist explanation of these disorders, they would say that the values changed and what was once a disvalued, hence deviant behavior, ceased to be with the shift in values. According to Murphy, constructivism does not have theoretical resources to correctly adjudicate these cases because these diagnoses were "rubbish all along, and scientific rubbish to boot, and we need a view that says that"(2006, p. 27). Thus, constructivism ignores an "essential fact about our everyday thought on mental illness, which is the importance of causal explanation"(2006, p. 28). In other words, Murphy argues that what is sought for is a "reason to regard the process underlying deviant behavior as themselves abnormal if the behavior is to strike us as symptomatic of mental illness." (Murphy, 2006, p. 29). He thinks that what is needed to consider some behaviors or processes as pathological rather than mere deviations from social norms, a "right sort of causal history"(Murphy, 2006, p. 29) meaning that there has to be some explanation as to what went wrong in pathological conditions. While constructivists can hardly come up with such solutions, objectivists offer an important element which Murphy favors by asking "what the right causal antecedents are."(Murphy, 2006, p. 29). Murphy goes on to defend objectivist revisionism throughout his book.

Murphy's taxonomy of positions is of great help in navigating the waters of the both the traditional and the meta-debate. While Murphy's attacks on constructivism are pretty much standard and can be encountered frequently in the traditional debate, I think Murphy neglects an important family of concepts traditionally found on the normative side. As Tim Thornton writes:

> Some philosophers and psychiatrists (to whom I will refer to as 'values in theorists' or 'value theorists' for short) argue that at the heart of the idea of illness is something that is either bad for the sufferer, or is a deviation from a social or moral norm. (Thornton, 2007, p. 13)

With the disjunction in the citation above, Thornton emphasises two horns of the normativist, or value-laden concepts. They either represent that which is bad for the sufferer (hinting at harm-like concepts[17]) or is a deviation from a social norm, which is how Murphy thinks of constructivism. Murphy considers and argues against just the latter while not really discussing the former. Thus, his arguments are directed at a general version of constructivism while there are peculiarities which his analysis fails to consider. Granted, harm-like criteria are not necessarily in opposition to Murphy's account. As we have seen in his analysis of Wakefield whom he considers conservative objectivist, harm is a separate criteria and dysfunction is not veiled nor clouded with harm's normativity. Harm-like concepts would be in stark opposition to Murphy's account if we considered them as a necessary element of disorder, if we thought that an objective/scientific element is not separable from the harm-like elements. In my work I claim none of this.

The lack of Murphy's attention to concepts that imply illnesses are 'bad for the sufferer' is somewhat understandable seeing as Murphy's project is for the most part centered on objectivism of the revisionary kind, considering his opponent in this the conservativist objectivist. He briefly and matter-of-factly eliminates constructivism in the beginning of his book while focusing mostly on the objectivist, scientific side.

I think Murphy's rendition of the debates can be enriched and supplemented by putting at a forefront an important set of concepts. These concepts refer to the

[17] Harm-like are concepts such as suffering, well-being, harm, welfare, flourishing etc. In the following I specify and articulate them more elegantly in the form of compassionatism.

state of the individual, usually from their own perspective and include concepts like harm, unluckiness, well-being, happiness. While in their nature mostly normative, pace Murphy, they do not necessarily subscribe to the view of mental disorder as social deviance. Most often they have been lumped together with concepts that reflect 'social deviance' since they are normative as well. However, they need not be relative to the society's standards and do not urge for social conformity. I call the position that considers these concepts *compassionatism* in philosophy of psychiatry. They address the state of the individual, focusing on elements of their own psychology and subjective experience. I emphasize these concepts and bring them to the forefront while also integrating them into Murphy's taxonomy of positions.

### 5.5 Compassionatism

Compassionatism is a view that emphasizes the first person's, the patient's perspective in psychiatry. These are concepts which instantiate the individual's experience, from their own point of view. Compassionatism is the embodiment of one of the three pillars of good psychiatric practice.

When it comes to compassionatism we can postulate two variants– conservative compassionatism and revisionist compassionatism. It is important to emphasise that this is an addition to Murphy's taxonomy. It does not go against his analysis of objectivism nor does it claim that the notion of disorder is essentially normative. Its function is to put to the forefront important elements of psychiatric interest which I feel weren't paid enough attention in the debate, as well as in Murphy's taxonomy. The pretense for making these concepts transparent is not to argue against any of the positions of the traditional debate, i.e. neither naturalist nor normativists. With regards to that, I remain agnostic. I only argue that concepts encompassed by compassionatism are and should be important in our psychiatric considerations. I argue that they are ineliminable criteria of psychiatric interest and represent one of the pillars of good psychiatric practice.

### *5.5.1* *Conservative compassionatism*

Similar to conservativism I have outlined earlier compassionate conservativism takes one of the concepts like harm, suffering, well-being and maps out how the concepts are used. It is a descriptive project that uses conceptual analysis to ascertain the necessary and sufficient conditions of the application of the concept.

We can interpret Cooper's (2020) analysis of harm in psychiatry as a recent example of conservative compassionatism. According to Cooper, the concept of harm is undergoing a slight conceptual shift in psychiatry. This is especially relevant for descriptive projects using conceptual analysis:

> For philosophers who aim to produce a descriptive analysis of the concept of disorder it matters that the consensus that harm is required for disorder is breaking down. We used to say that harm is required for a condition to be a disorder, but now some people will refer to harmless dysfunctions as disorders. If our linguistic community starts to employ terms differently (especially if experts start to employ a term differently), then conceptual analyses will also have to change. (R. V. Cooper, 2020, p. 148)

Thus, in order to analyze harm descriptively using conceptual analysis one has to be vary of the conceptual shift. Cooper is describing how the concept of harm changes without taking a normative stance on it. Instead, she methodologically proposes a solution that might aid the conceptual analysis by belt-and-bracing certain considerations that seem to be robust. The crux of conservativist compassionatism is describing or mapping how some of the concepts such as well-being, harm, suffering etc., are being used or being thought of by the stakeholders in the debate. It is doing

conceptual analysis on such concepts that is curtailed by what the stakeholders, both lay and professional think the concept represents.

However, I am not interested in that project. I think that to successfully analyze harm and to come up with novel and creative solutions of approaching it, one has to be a revisionist. This does not mean that there is not an element of conceptual analysis to the project but it makes a part, not the bulk of it. In the following I outline revisionist compassionatism.

### ***5.5.2*** *Revisionist Compassionatism*

Both conservative and revisionist compassionists deal with concepts that take into account the first person's perspective, from their own point of view, and pay serious mind to the good and interests of the person. The concepts that reflect this care are suffering, harm, well-being, welfare, unluckiness and any such concept that is evocative of this perspective.

Conservative approaches aim to describe these concepts using conceptual analysis. They are interested in tracking these notions as the stakeholders in the debate view them. Conversely, revisionist compassionatism, as all revisionist positions, does not aim to describe but prescribe. Instead of asking what the concept *is*, the revisionist compassionatism consists in thinking about how these concepts such as well-being, harm, suffering, welfare and so on *should be*.

Now, the question arises as to what the criterion is according to which these concepts are judged and designed. Since we are not merely describing these concepts but prescribing their use, role and constitution, what should be our guiding principles in constructing them? In the first chapter I have outlined the three pillars of (good) psychiatric practice – psychiatric expertise, scientific evidence, and patient's perspective. The role of these concepts is and should belong to the domain of patient's perspective. Since this is the role of these concepts, it hints at where the first criterion lies. The first principle of revisionist compassionatism is to fully and most

faithfully capture the individual, patient perspective – the patient's unique and complex point of view. It is to put considerable efforts into figuring out how the condition affects their life. This mostly transpires in the psychiatrist-patient relationship where the psychiatrists based on interviews and psychological assessments determines the state of the patient. Diagnosis is a collaborative effort where each side plays a part in coming to understand the other. However, adopting compassionatism means approaching the patients with openness, compassion and interest in their story – their life from their own point of view.

However, this principle is constrained and further specified by the second principle. The second principle takes into account other considerations in practice that are represented by the two other pillars of psychiatric practice – scientific evidence and psychiatric expertise. The second principle thus states that the patient's perspective should be taken and interpreted in a way that is intelligible and commensurable with other parts of psychiatric practice. For example, it may be the case that the patient thinks a work of literary fiction like Sylvia Plath's *The Bell Jar* is a great depiction of mental disorder and that it faithfully describes their point of view. However, such references are inadmissible when it comes to diagnosis. Rather, the psychiatrist must try to see what exactly speaks to the patient in such a work and to translate their views and experiences into psychiatric terminology and symptomatology. In other words, the first principle advocates for listening, encouraging and eliciting the patient to share his worldview, perspective and issues in the context of their life. The second amounts to interpreting and translating the patient's perspective according to the rules and theory of psychiatric practice. Thus, these two principles should be adhered to in unison during the process of diagnosis.

There are at least two reasons why we should establish revisionist compassionatism as a cornerstone of psychiatric practice. In the chapter on depathologizing homosexuality I have emphasized the importance and value of using

concepts such as harm to capture the patient's perspective. The unprecedented move from 1970s in depathologizing homosexuality was a monumental step towards the end of psychiatric abuse and misuse of psychiatric authority. The history of psychiatry is notorious for medicalizing conditions based on the value systems of their time with little regard for the view of the persons 'suffering' from them and with little to no scientific validity. While the diagnosed had a huge stake in their diagnosis, as all people who are diagnosed have, they had almost no say as to how the 'condition' affects them and whether it is actually a problem for them. 'Fictive' disorders like masturbation, drapetomania, hysteria and many more are a testament to this psychiatric blind spot and to the selective deafness and carelessness exhibited towards the experiences of the individual human condition in psychiatry. The notion of harm entering the mainstream psychiatric practice in the 1970s signifies and should signify the end of such misuse of psychiatric authority. This is why we need to hold onto harm and/or similar notions in psychiatry that encapsulate the perspectives of the patients, which is a view I have dubbed compassionatism. A lot of mistakes and lives ruined can be avoided by thinking in compassionate terms about how condition affect the receivers of health care. And hopefully, by employing such concepts, the misuse of psychiatric authority can be prevented in the future.

Additionally, compassionist concepts do not only have a role of preventing negative outcomes but offer novel insights and understandings of psychiatric conditions. In the process of depathologizing homosexuality we have not only been warned as to the dangerous of psychiatric misdiagnosis but have been offered better, more fuller and more faithful understanding of homosexuality as a condition. Without careful investigations and analysis of personal perspective this kind of insight would never have been possible. Thus, the role of the compassionate concepts is twofold – they prevent dangerous misuse as well as offer great insight into intricacies of conditions. Both can undoubtedly help shape the diagnosis, treatment as well as

prevention of mental disorders. Thus, I believe compassionatist concept such as harm, well-being, welfare and so on should be present, acknowledged and paid considerable attention to in the psychiatric setting.

Besides compassionatism, I believe revisionist attitudes should be the permanent atmosphere inside psychiatry. A lesson from wrongful pathologization of conditions from history of psychiatry not only attests to the importance of first person's perspective (which compassionist concepts embody) but also to the careful skepticism of our current practices. In studying history of psychiatry, the leitmotif that comes up again and again is that we often get things wrong. For example, throughout history, hysteria was thought of as a malady that affects exclusively women because of their female organs. This 'condition' was thought to affect emotional, dominant or disobedient women, being a "medical explanation for everything that men found mysterious or unmanageable in women" (*The History of Hysteria*, n.d., para. 9). The cure for such a 'condition' was gynecological and uterine massages and stimulation (which of course women could not do themselves) until the point of hysterical paroxysm (what we now call an orgasm). In severe cases the treatment was hysterectomy - the removal of female reproductive organs. To someone of this day and age, this might sound as an absurd remnant of the history of medicine. However, hysteria has a long history dating back even to Ancient Egypt and it was not long ago (in 1950s) that hysteria was excluded from the APA's classification of diseases. Homosexuality was depathologized not long ago as well. Homosexuals still alive today might have had the unfortunate experience of living through the time when it was still considered a mental disorder. These and many more unrightful depathologizations are evocative of the historical fact that psychiatry has a long track record of getting things wrong. It would seem as scientifically callous and epistemologically irresponsible to claim that all of the classifications and diagnoses which psychiatry is working with today are just right. Chances are that in years to

come with scientific advancements and careful investigation of psychiatric conditions we might witness breakthroughs which would discard the classifications and diagnoses that we have today. That is why holding to revisionist attitudes – not only describing how psychiatric conditions are configured, but also thinking long and hard about how they should be configured, might be the key to mitigating and preventing the mistakes repeatedly encountered in the history of psychiatry.

These are the reasons I believe it makes sense to hold onto revisionist compassionatism in psychiatry and to make it a salient position rightfully deserving of investigative efforts. I dedicate the remainder of my thesis and employ such efforts in investigating of one such compassionatist concept – the notion of harm in psychiatry.

## 5.6 Conclusion

This chapter deals with the meta-debate on mental disorders. I adopt Murphy's taxonomy of the meta-debate that includes objectivism and constructivism, both of which can each be either conservative or revisionist. I outline the upsides and downsides of all of the permutations of the positions. Murphy is a representative of the revisionist objectivism and argues against constructivism. However, I believe Murphy does not pay proper attention to the harm criterion, usually attributed to constructivism in the traditional debate. Instead of making a case for or against any of the strands Murphy proposes, I offer an addition to Murphy's taxonomy in the form of compassionatism. Compassionatism serves to highlight concepts that take into account the patient's perspective, such as well-being, suffering, harm, and welfare. I defend a revisionist variant of compassionatism and propose principles which compassionist concepts should adhere to. First principle is that they should as fully and as faithfully capture the patient's perspective. The second is that these concepts should do so following the principles, guidelines and concepts of psychiatric practice. In the following chapters I go towards building the account of harm that satisfies these criteria.

# 6. Towards an Account of Harm in Psychiatry

## 6.1 Introduction

In the previous chapter I outlined and defended revisionist compassionatism. It is the view that concepts such as harm, well-being, welfare, suffering and so on, which I have grouped under the heading of compassionatism, are and should be salient in psychiatric practice and are deserving of its rightful place. I argued that compassionatist concepts embody the patient's perspective which I have argued earlier is one of the three pillars of good psychiatric practice. There are two principles to adhere to in constructing of revisionist compassionatism – the concept should as fully and faithfully capture the patient's perspective and it should do so in a way that is intelligible to the psychiatric expertise and practice. I dedicate this chapter to setting the groundwork for building the account of harm. First, I pay mind to the methodology of the project and decide on using philosophical explication inspired by Rudolf Carnap and enriched by Peter Strawson. I then postulate the notion of harm in bodily medicine that consists in the organism's ability to achieve and maintain homeostasis. I translate this account of harm to psychiatry.

## 6.2 Methodology

Since I propose the revisionist modality of compassionatism it means that we are not only interested in how the concept actually is, the conceptual analysis of it, but we are thinking about how these concepts should be construed. Since it is not a descriptive, but a creative project, there are two criteria or principles that should

guide our inquiry. The concept should as fully and as faithfully capture the patient perspective. And it should do so in a way that is understandable and intelligible to the psychiatric practice. These are the principles according to which we should build as well as evaluate the success of our compassionist concepts.

One such concept – the notion of harm, is the focus of this thesis. The ultimate goal is to come to a satisfying solution about how harm in psychiatry should be conceptualized. We are not looking to capture the commonsensical concept of harm from all of the stakeholders in the debate. What we are looking for is concept that would serve a particular purpose in a specified domain – psychiatry and medical practice. Thus, this should be reflected in the methodology of our project.

The methodology I employ in coming to the concept of harm is philosophical explication. We have come across philosophical explication in the previous chapter as it is a method often favored by revisionist projects in general. The pioneers of philosophical explication are Carnap (1962) and Quine (2013), but it has recently regained its popularity in contemporary philosophy as well (Brun, 2016; Dutilh Novaes, 2018; Schwartz, 2007b).

Philosophical explication is a method by which we take an inexact concept and shape it according to our needs or wants into a concept that serves a particular purpose. As Carnap (1962, p. 3) explains "The task of explication consists in transforming a given more or less inexact concept into an exact one or, rather, in replacing the first by the second.". The term used for the first concept, the explicated-to-be one, is explicandum while the latter, the concept we get through explication is called explicatum (Carnap, 1962, p. 3). The explicandum may be derived from natural language or be some scientific concept that requires further elaboration. The explicatum is our end-result of the process of explication.

To illustrate the method of philosophical explication Carnap uses the example of 'Fish'. Our prescientific meaning behind the concept 'Fish' mostly included

'creatures living in the water' such as cold-blooded vertebrates with gills but also seals, whales, dolphins and other various kinds of sea creatures. While in terms of everyday language we could generally use 'Fish' to refer to many of these creatures without much problem or confusion, in terms of scientific taxonomies and advances in zoology the concept was problematic in that it was too broad, including all kinds of sea creatures of very different biology and origin. To resolve this, the term 'Piscis' was brought into scientific discourse to signify just those animals that live in the water, are cold-blooded vertebrates and have gills throughout their lives (Carnap, 1962, p. 6). Thus, mammals like whales and dolphins and many more were excluded from the scientific concept of 'Piscis'. We ended up with the concept of 'Piscis' which is narrower and more specific than 'Fish'. Additionally, it fits better into the scientific context and its precision and elegance is easily implemented in our theories about the natural world. Applying Carnapian terminology to this example, the term 'Fish' is explicandum, i.e. the notion to-be-explicated and 'Piscis' is explicatum, i.e. the end result of explication.

Although philosophical explication is a normative project, distinct from conceptual analysis which is a descriptive one, conceptual analysis still has its place in the process of explication. Even more so, conceptual analysis is its first step. In order to explicate a particular concept we have to have an idea of what that concept represents, however vague or inadequate that concept may be. The concept first has to be clear enough so that we can begin our process of explication. If we have no grasp or notion of what the original concept entails, we are tapping in the dark as to how it should be explicated. Explication essentially entails specifying the concept according to some desiderata so how can we specify a concept if we do not have anything in the first place *to* specify? This step is what Carnap (1971, p. 4) calls 'formulating the problem':

> There is a temptation to think that, since the explicandum cannot be given in exact terms anyway, it does not matter much how we formulate the problem. But this would be quite wrong. On the contrary, since even in the best case we cannot reach full exactness, we must, in order to prevent the discussion of the problem from becoming entirely futile, do all we can to make at least practically clear what is meant as the explicandum.

Luckily, in looking for our concept of harm we do not have to start from scratch. There are two options we can resort to in formulating the explicandum of the concept of harm. One is to take one of the notions of harm already present in the debate on mental disorders which I have outlined in the second chapter of this work. That, however, is a bit problematic. As I have indicated the problem with the existing conceptions of harm is that they are either problematic and not really satisfactory, as in the case of distress or disability (see 2nd chapter), or that they are vague and underspecified. Another route is to go back to medical practice and to try to find some other notion of harm on which to base our process of explication. We can start from the notion of harm that I argue is already embedded in the medical practice. I believe it is the rationale behind numerous conditions that we consider as disorders in bodily medicine and that it may be present in psychiatry as well.

I propose that in formulating the problem, which is the first step in the process of explication, we resort to a particular and distinct way of doing conceptual analysis proposed by Peter Strawson. In his book *Analysis and Metaphysics* (Strawson, 1992) he reflects on the aim and method of analytic philosophy. He starts by using an analogy. Strawson tells the story of presenting the first Castilian grammar to the Queen Isabella of Castile. After being instructed as to what grammar is and what it entails, she inquired what would be the use of such a project. In her mind, those who spoke Castilian had no use of grammar since they already followed these grammatical

rules. Granted, it was a time unlike ours where you would not often see someone not speaking your native language. Strawson explains further:

> The grammar did not set the standard of correctness for the sentences they spoke; on the contrary, it was the sentences they spoke that set the standard of correctness for the grammar. (Strawson, 1992, p. 5)

Speakers of Castilian already used the principles and the rules of grammar since they started acquiring the language, even as soon as after the birth. Providing them with the grammar would seem as an afterthought. In one sense, they already knew the grammar of their language since their language use flawlessly reflected it. However, Strawson points out that there is also a sense in which the native speakers did not know the grammar of their language. After all, language use of native speakers does not entail learning the formal aspects of the language such as grammar. Rather, the speakers of Castilian, as well as of any other language acquired the rules of their native language implicitly. This means that they acquired the language through immersion in the language context, through being surrounded and communicated with in their native tongue, through learning from others who spoke and by observing all of the language elements in real time. If they were asked to state the rules and principles they adhered to in their use they probably would not be able to do that. Strawson concludes that "the general moral that being able to do something in this case speak grammatically is very different from being able to say how it's done; and that it by no means implies the latter." (Strawson, 1992, p. 6). Strawson transfers his analogy of grammar to philosophy. Strawson (Strawson, 1992, p. 6) argues that

> rational human beings, capable of developed thinking, must have an implicit mastery of more than grammars; or, rather, their implicit mastery

> of their grammars is intertwined with an implicit mastery of all the concepts, all the general ideas which find expression in their speech, which they operate with in their thought.

Thus, he uses the grammar analogy to make a point about our conceptual apparatuses arguing that we "we operate with an enormously rich, complicated, and refined conceptual equipment" (Strawson, 1992, p. 6). However, similar to grammar, we have not been and probably could not have been taught these intricate meanings and relations between the concepts by being explicitly instructed as to their theory or principles guiding it. Rather, we use and have understanding of all kinds of concepts implicitly, meaning we might possess and work with a given concept without being able to know or explain what it exactly entails. Thus, even though we 'know' some concept we may not be able to explicate it.

Thus, Strawson (1992, p. 7) argues "Just as we may have a working mastery of the grammar of our native language, so we have a working mastery of this conceptual equipment.". And just as the grammarian aspires to map out the rules, principles and structures of a language, "so the philosopher labours to produce a systematic account of the general conceptual structure of which our daily practice shows us to have a tacit and unconscious mastery." (Strawson, 1992, p. 7). According to Strawson, the goal of philosophy is to make explicit the rules and principles that govern our conceptual apparatuses. It is to map the conceptual landscapes we operate on and to make concepts and relations between concepts explicitly known.

Meanwhile, Strawson uses the analogy between grammar to make a further point. He argues that the this 'traditional' method of conceptual analysis persists in contemporary analytic philosophy and is a guiding model of much of it. Philosophers are breaking down a concept into elements that the concept has to possess to be that particular concept. In other words, they are searching for necessary and sufficient

conditions for the application of a particular concept which is nothing if not a prototypical activity of analytic philosophers. Therefore, when it comes to the method of conceptual analysis the bar is set relatively high meaning that both any possible counterexamples or circularity have to be avoided. This sets many of the philosophical projects back as it is very difficult to provide definitions that avoid both of these problems. For example, consider the definition of knowledge as justified true belief[18]. After Gettier's (1963) counterexamples to this traditional tripartite definition of knowledge, it seems that the discussion has been going around in circles. The standards of the discussion seem to be too high as one counterexample was a knock-down objection to the theory and the tripartite definition of knowledge was not able to recover ever since. However, whether these intricate and elaborate systems that try to provide alternatives in wake of Gettier's counterexamples actually bring us closer to knowing and understanding the concept of knowledge is debatable. As a result, many authors turned to novel ways and ideas of approaching the concept of knowledge. Similar climate to the discussion on knowledge is under way in philosophy of psychiatry as well. Since the traditional discussion has been heavily objected to (Ereshefsky, 2009; Murphy, 2006; Schwartz, 2007a) (see 3rd chapter), it seems like the right time to introduce novel ways of understanding and dealing with the notions in philosophy of psychiatry.

Thus, Strawson (1992, p. 19) introduces a new model that is purported to be "more realistic and more fertile" than the standard view of conceptual analysis in terms of necessary and sufficient conditions. He urges us to

> abandon the notion of perfect simplicity in concepts; let us abandon even the notion that analysis must always be in the direction of greater

[18] The definition of knowledge as justified, true belief dates back to Plato (Parikh & Renero, 2017) and has been the accepted view until the 20th century.

> simplicity. Let us imagine, instead, the model of an elaborate network, a system, of connected items, concepts, such that the function of each item, each concept, could, from the philosophical point of view, be properly understood only by grasping its connections with the others, its place in the system perhaps better still, the picture of a set of interlocking systems of such a kind. (Strawson, 1992, p. 19)

He dubs this account of analysis 'connectivist' as its aim is to bring elucidation and understanding by connecting concepts in a coherent and meaningful network. In such system the concepts are illuminated by referring to other concepts which may often be simpler, or we could even call them more primitive. The basic idea is that the concepts are illuminated through connection with other concepts. This connectivist idea of analysis puts less strain on our definitions and makes the process of concept analysis more palatable as we are no longer trying to align to very high standards of defining concepts in terms of necessary and sufficient conditions. Strawson calls this traditional project of conceptual analysis in terms of necessary and sufficient conditions reductivism as we are reducing the concept to their (necessary and sufficient) elements. Reductivism is juxtaposed with what Strawson dubs connectivism. Connectivism is a more modest approach, that of elucidation and illumination of concepts through their connection to other concepts. Since I have mentioned before, the more traditional, reductivist approach has many adversaries (see the 3rd chapter) so connectivism can serve as a breath of fresh air in our methodology.

In constructing the account of harm, the first step is to adopt this Strawson's idea that one of the jobs of philosophers is to make an implicit concept explicit. Additionally, in using the Strawsonian particular flavor of conceptual analysis we do not need to look for necessary and sufficient conditions but to use concepts that

illuminate and deepen our understanding of harm. I try to delineate the notion of harm that is used in medicine and I cover this in the following section on harm in bodily medicine.

While Strawson's ideas might work for philosophical concepts in general, here we are dealing with specific and rather technical concepts that are ingrained in the field of medicine and psychiatry. Strawson's ideas are just the first step in our method of philosophical explication. They are used for formulating the problem, getting the idea of our explicandum. Let us see the further elements that can be expected in the process of explication by returning to Carnap's view of explication.

Carnap offers four principles, requirements that the end concept must satisfy in coming to the explicatum. These requirements are: "(1) similarity to the explicandum, (2) exactness, (3) fruitfulness, (4) simplicity (Carnap, 1962, p. 5). Carnap elaborates on each of these requirements:

> I. The explicatum is to be similar to the explicandum in such a way that, in most cases in which the explicandum has so far been used, the explicatum can be used; however, close similarity is not required, and considerable differences are permitted.
>
> 2. The characterization of the explicatum, that is, the rules of its use (for instance, in the form of a definition), is to be given in an exact form, so as to introduce the explicatum into a well-connected system of scientific concepts.
>
> 3. The explicatum is to be a fruitful concept, that is, useful for the formulation of many universal statements (empirical laws in the case of a nonlogical concept, logical theorems in the case of a logical concept).

> 4. The explicatum should be as simple as possible; this means as simple as the more important requirements (1), (2), and (3) permit. (Carnap, 1962, p. 7)

We see that not all requirements hold equal value and importance. Carnap gives primacy to fruitfulness. The reason behind this could be that the criterion of fruitfulness carries some of the most important aims of explication – easing the concepts' integration into the scientific context and formulation of natural laws. Exactness is almost as important as fruitfulness because it participates in the execution and integration of fruitfulness into the scientific domain. Then comes similarity which should be constrained by fruitfulness and exactness. While the concept should entail at least the remnants of a given, original concept this comes second to the pragmatic utility of the explicated concept. Lastly, there is simplicity. More generally, simplicity has been one of recurrent criteria in philosophy of science when it comes to virtues that theories should uphold. As Baker writes "There is a widespread philosophical presumption that simplicity is a theoretical virtue." which has been defended throughout the history of Western philosophy – from the Ancients to the Contemporaries (2004, sec. 1. Introduction). While it is generally viewed as better that a theory makes as few ontological commitments and includes as few theoretical elements as it needs, Carnap considers simplicity to be one of the least important criteria, coming last to all others. In other words, fruitfulness, exactness and similarity should not be sacrificed in the name of simplicity.

Let us then assess Carnap's example of 'Fish' and 'Piscis' to illustrate the ideas behind his requirements. When it comes to fruitfulness it seems that 'Piscis' is more fruitful than 'Fish' in that it is more appropriate for scientific contexts and has merit when it comes to formulation of theories and connection to the

biological body of work. It is more precise, 'cuts' the concept for its useful properties while excluding the other less helpful properties. 'Piscis' is a unitary concept for a group of similar phenomena worthy of being characterised as such. 'Piscis' also is a more exact term, more narrower and has a clearly articulated extension. When it comes to similarity it is similar to the concept of 'Fish' in the important ways since we can observe the connection between the two concept and its similarity in meaning. Lastly, 'Piscis' is also simpler than the concept of 'Fish' as it is less theoretically burdensome and includes fewer but more specific elements.

Thus, when it comes to Carnap's ideas on methodology, we can emphasize two important aspects. Explication *a la* Carnap is a normative project. This means that we are not only asking ourselves how the concept is used by describing its various instances of usage. Rather, we are asking ourselves how the concept should be used to correspond to our theoretical and scientific interests. Explication is a process whereby we take a less exact concept and configure it to fit some prior established desiderata. This process, however, does not mean that 'anything goes' as Carnap suggests four requirements that the explicatum has to satisfy – similarity, exactness, fruitfulness and simplicity.

Combining Strawson's and Carnap ideas we end up with the following methodological recommendations in addressing the concept of harm. The first step in the process of explication is formulating the problem. This means we need a satisfying grasp on explicandum in order to get to the explicatum. I use Strawson's ideas on conceptual analysis to get there. According to him, philosophers' job is to make explicit the concepts we implicitly acquire and apply in our everyday lives. Additionally, Strawson proposes connectivism, an account according to which instead of searching for the necessary and sufficient

conditions, we provide meaning and understanding of the concepts by connecting them into a conceptual system and seeing the relations between the concepts. This is the first step in our analysis and serves as a way to formulate a problem, the concept we are starting from which Carnap calls explicandum.

After we have done that, Carnap's subsequent steps in the method of philosophical explication come into play. Its aim is to take a given concept and specify it, tailor it to our needs, so as it can be better integrated in our scientific practice. In doing this, we should be mindful of the principles of similarity, exactness, fruitfulness and simplicity in devising our concept, i.e. explicatum.

In the remainder of this chapter I follow the above laid out methodology. To simplify and emphasize the structure I summarize the methodological steps:

1. Elucidation of the concept of harm in medicine – making explicit what has been taken as an implicit when it comes to harm
2. Asking ourselves the question of what the concept of harm should look like in psychiatry. Gathering up desiderata and constructing a concept of harm following the ideas of philosophical explication.
3. Checking whether the requirements of similarity, fruitfulness, exactness and simplicity correspond to our derived concept of harm

### 6.3 Harm in (Bodily) Medicine

Regarding harm in bodily medicine, I contend that, even though not explicit, much of what we consider under contemporary medicine operates under some kind of idea of harm. As I have emphasized in the previously, our task as philosophers is to make explicit notions, ideas, concepts that are implicit in our conceptual apparatuses. Many physicians might be using the notion of harm unbeknownst to them. It is instrumental

for our exploration of harm to bring these considerations to the forefront. I argue that there is a tentative and implicit assumption of harm that underlies numerous instances of medical practice. In devising an account of harm for psychiatry I deem it as important to start from the ideas of harm in medical practice in general in bodily medicine.

Medicine is a broad field of study including many disciplines and areas of inquiry. Since we are dealing with the notion of disorder or disease, a field that is central to our exploration is pathology – "scientific study of changes in the structure and function of the body in disease." (Mohan, 2015, p. 1). As treating disease is usually seen as the aim of medicine, and health is often defined as an "absence of the disease"(Boorse, 1977), the knowledge of pathology is essential for scientific and professional medical practice. As Mohan argues:

> knowledge and understanding of pathology is essential for all would-be doctors, as well as general medical practitioners and specialists because unless they have knowledge and understanding of the language in the form of pathology laboratory reports, they would not be able to institute appropriate treatment or suggest preventive measures to the patient. (Mohan, 2015, p. 1)

To seek some universal rule or notion that is present and common to all fields of medicine or that may underline the field of pathology might at first seem as a quite ambitious project. However, there are theoretical considerations that are at play in medicine which can be considered as common to many medical subfields. One might delineate them as meta-medical theoretical setups that are in the background of the way medicine is researched and practiced. I argue that this can be achieved by using the concept of homeostasis. Mohan hints at this by using the concept of homeostasis

in his specification of pathology "Pathophysiology, thus, includes study of disordered function (i.e. physiological changes) and breakdown of homeostasis in diseases (i.e. biochemical changes)." (Mohan, 2015, p. 1). The work of George Billman (2010, 2013, 2020) is especially instructive in this regard as he considers homeostasis one of the central concepts in medicine. In his paper *The great challenge to physiology: to integrate function from molecules to man* (Billman, 2010) he outlines what he considers to be one of the greatest challenges in medicine. He introduces the dichotomy between holism and reductionism. Reductionism "attempts to explain the nature of complex phenomena by reducing them to a set of ever smaller and simpler components", it is "the view that the whole is merely the sum of its parts"(Billman, 2010, p. 1). While the propensity for thinking about matter or organisms in terms of their smaller parts has been present since Ancient Greece with trying to identify the ultimate underlying substance (*arche*) and with Lucretian and Epicurean atomism, in contemporary times it has become the dominant approach to medicine and science in general (Billman, 2010, p. 1). As Billman states "There can be no denying the power of this approach"(Billman, 2010, p. 1), especially since the discovery of genes and DNA sequencing. The great challenge for physiology then "is to integrate and to translate this deluge of information obtained *in vitro* into a coherent understanding of function *in vivo*" (Billman, 2010, p. 2). In other words, it is to transfer the knowledge and understanding at the micro level into the macro level, from the mechanism to the organism. Our faith should not only be put in understanding particular mechanisms, cells or molecules but how these mechanisms and cells work together in an organism.

Thus, instead of only taking the reductionist approach, Billman advocates for a holistic, homeostatic approach[19]. He argues that

[19] From this point forward I will use the term 'homeostatic approach' as a particular flavor of holistic approaches that I argue for.

> A greater emphasis must be placed on the traditional integrated and more holistic approaches developed by the scientists who gave birth to physiology as an intellectual discipline. In other words, it is time for physiologists to return our roots. (Billman, 2020, p. 2)

He argues for and concentrates on the concept of homeostasis which he defines as "a self-regulating process by which biological systems maintain stability while adjusting to changing external conditions." (Billman, 2020, p. 2). According to Billman (2010, p. 2), homeostasis is how an organism can sustain its internal conditions while faced with external circumstances. The idea behind homeostasis is that we are self-regulating organisms composed of elements and mechanisms that work together to achieve equilibrium. This idea is by no means novel as it dates back even to Ancient Greece. Hippocrates was one of the authors who advocated for the "nature's helping hand in the healing process (*vis medicatrix naturae*)" (Billman, 2020, p. 1; Hall, 1975) and emphasized the body's ability to heal itself. The concept of homeostasis captures important thoughts on medicine. It was argued that the physician is there to aid nature in its course, to help mechanisms in an organism work in establishing and restoring balance. Self-regulation has been seen as the driving force of achieving health, with homeostatic approaches explaining disease in terms of body dysregulation and health as body regulation. As Billman notes, "The disruption of homeostatic mechanisms is what leads to disease, and effective therapy must be directed toward reestablishing these homeostatic conditions, working with rather than against nature."(Billman, 2020, p. 2). The organism strives to gain and maintain internal stability in the wake of outside forces. For example, outside we are faced with elements such as wind, rain, sun and so on. Our bodies work to keep for example our temperature, glucose level, oxygen level and all other functions of our body stable. This means that the whole system of our bodies works to compensate for the outside

factors. When it is cold outside our bodies work to keep our temperature stable and our bodies warm – our blood vessels constrict so to reduce the volume of blood that flows in order that our blood stays warm, the blood is sent to the vital organs rather than extremities and so on to protect our organism. When we are hot, our bodies produce sweat whose function is cooling our bodies off. There are numerous mechanisms like these that are in place so that our bodies and their functions remain stable in light of outside conditions.

Thus, achievement and preservation of homeostasis seem to be what organisms strive for. The world both natural and social is a rough place. We are constantly surrounded by external influences that have the potential to take a toll on our organism. But because we are of this world and borne into it, our organisms have systems in place that mitigate these external factors since the day we were conceived and since the beginning of our evolutionary history. This is how we fend off diseases and how are bodies are able to interact with the environment but remain relatively unscathed by it. When either achievement or preservation of homeostasis of the organism is compromised to the point where the organism cannot fend for itself or achieve homeostasis on its own, then we have a problem. This is when it makes sense to talk about harm in bodily medicine. We are harmed when our organisms cannot achieve or maintain homeostasis by using its own resources. Specifically, we are harmed when our organisms cannot achieve and maintain homeostasis internally while being subjected to external factors and environment. In other words, when an organism cannot 'deal' with some problem on its own – when its ability for achieving and preserving homeostasis is brought into question, then it is harmed. This is when we do (and should) seek medical assistance. Following the same rationale, medicine thus helps our organisms to achieve and maintain homeostasis. It works together with our organisms to fight off and neutralize looming threats. I argue that these ideas are deeply ingrained in medicine and have been, more implicitly than explicitly, the

underlying rationale of medical practice. The idea of homeostasis is by no means new as the concept has been ever-present in the history of medicine – going from ancient times to the present (Billman, 2020) with reductionism only gaining traction in recent times. To illustrate these theoretical underpinnings of medicine let us consider few examples.

Take the example of a broken leg. There are several treatments for such condition depending on the state in which the leg is. If the bone is in place the treatment is immobilization which includes keeping the leg in the fixed position so that it can heal, and future injuries can be prevented. Most common type of immobilization is putting on a cast or a splint that fixates the leg's position. If the bone is out of place it first needs to be adjusted which is a process called 'reduction' (*Broken Leg*, 2017). After that is done, the leg is treated by putting on a fixature to keep it in the same position so that healing can take place. Sometimes, there are complications and the broken leg requires surgery by which they might fixate the bones by using screws or some other method to recreate the closest anatomical formation of its previous state. After that the process is the same – putting a kind of cast or splint to keep the leg in the same position to facilitate healing. We see here that with regards to leg fracture, medicine relies on the body to heal itself and it just 'aids nature' as I have characterized it before. Broken leg represents a dysregulation in the system – the body might need help in healing the condition either by immobilization, bone adjustment or some other kind of medical intervention. However, after these things are in place it is on the body to heal the injury and re-establish homeostasis. It seems then that the organism is harmed when it is unable to restore homeostasis on its own. When the bone in the leg is misplaced, it is almost impossible for the organism itself to restore it to the original position internally. However, once that is done medical intervention relies on the body's healing system. We do not have a 'cure' for broken legs per se, we only facilitate healing by relying on the systems in our organisms and

by keeping the leg in the fixed position so that the healing can take place. Harm consists in the inability of an organism to respond through its internal resources to external challenges. In lay terms, the organism is harmed when it can't solve some 'problem' on its own. It is then when organisms need and can use the help from medicine.

The example of the broken leg is not isolated as the rest, or the most of pathological condition follow the same pattern. Case in point are bacterial infections. Bacteria are all around us and are an "essential component of the human microbiome, colonizing tissues including the gastrointestinal tract and skin." (Deusenbery et al., 2021, p. 695). Most of these bacteria are harmless and can even benefit us, for example, gut bacteria that help and facilitate digestion (Deusenbery et al., 2021). To investigate the way bacterial infections work to impede our health we need to explore bacterial pathogenesis which "refers both to the mechanism of infection and to the mechanism by which disease develops" (Peterson, 1996, sec. Introduction). Bacterial pathogenesis "is caused by a disruption to the normal coexistence of bacteria and host cells observed in healthy individuals." (Deusenbery et al., 2021, p. 695). Bacterial infections depend on the dynamic process between at least two variables – "the physiologic and immunologic condition of the host and on the virulence of the bacteria." (Peterson, 1996, sec. Host Susceptibility). Virulence is the "measure of pathogenicity of the organism", meaning that virulence is the degree to which the bacteria have the potential to pathologically affect the host. Virulence is

> related directly to the ability of the organism to cause disease despite host resistance mechanisms; it is affected by numerous variables such as the number of infecting bacteria, route of entry into the body, specific and nonspecific host defense mechanisms, and virulence factors of the bacterium. (Peterson, 1996, sec. Introduction)

In other words, whether an infectious disease develops depends on the body's defense system (the immune system) and the properties of bacteria attacking the organism. Our organisms are in constant negotiation between these two systems. Disease develops when the bacteria significantly compromise, overpower or bypass the defense mechanisms that are in place. In lay terms, this can happen if the organism's immune system is for some reason too weak or weakened, or if the influence of bacteria is too strong (significant increase in bacteria population, bacteria that are considerably pathological and so on). Our organisms are constantly subjected to this dynamic process between our internal mechanisms and external threats that are potentially harmful for us. Thus, for some organisms, namely infants (who hadn't had the chance to develop the immune system), older population (immune system weakens with age), and persons who suffers from some immunocompromising condition such as HIV, even the least pathological bacteria can cause disease. This happens because the immune system is too weak to deal with bacteria. Conversely, in healthy persons who have a normal functioning immune system, there are bacteria that can do significant damage and cause disease if the bacteria are so vicious and pathological that even a healthy system cannot deal with them.

The most common treatment for infectious diseases are antibiotics – "medicines that fight infections caused by bacteria in humans and animals by either killing the bacteria or making it difficult for the bacteria to grow and multiply." (CDC, 2022). Antibiotics work to help our system fight off the bacterial infection. Thus, when our organism cannot fight off bacteria on its own, namely when it cannot achieve homeostasis on its own, antibiotics are introduced to help the organism achieve homeostasis. Simplified, when our organism fails to deal with the problem of bacterial infection on its own, medicine helps by introducing resources to help deal with the infection. These resources attack bacteria so that our systems can easily fight them

off and control the growth of bacteria population. The organism is harmed then when our defense system is overpowered by the bacteria and that is when disease sets in. This is when we seek medical attention to help our organism fight the disease. Now, comparing this with the instance of leg fracture we observe significant similarities in rationale and theoretical set up of both conditions. In both cases we rely on the body to 'solve the problem'. Once the problem cannot be solved by the body, when the homeostasis cannot be achieved internally by the organism alone, then we seek medical attention to combat diseases.

Cancers are another instance of this 'failure of the organism to deal with the problem on its own' going in favour of the homeostatic view. There are different kinds of cancer but many of them seem to share a common mechanism– "Cancer arises when, within a single cell, multiple control systems malfunction." (Sompayrac, 2004, p. 3). These control systems are of two kinds: "systems that promote cell growth, and safeguard systems that protect against "irresponsible" growth." (Sompayrac, 2004, p. 3). The systems that promote cell growth determine under which conditions the cell grows in size and divides into two daughter cells which is a process called cell proliferation (Sompayrac, 2004). Cell proliferation is good, even majestic and miraculous as it is responsible for the fact that we go from babies to grown people. However, there are instances when cell proliferation is not controlled, and this is where tumours or unwanted growths can develop. This still does not necessitate cancer development since there are various instances of benign cell proliferation, for example skin tags which are not as aesthetically pleasing but are not life-threatening. Besides cell proliferation there are other systems responsible for guarding against cancer as they are in place to "protect against inappropriate cell proliferation" (Sompayrac, 2004, p. 7). These systems are of two types: "systems that help prevent mutation" and "systems that deal with mutations once they occur" (Sompayrac, 2004, p. 7). Each day there are mutated cells in our bodies which have the potential to cause

cancer, however, because we have these safeguarding mechanisms in place they neutralize most if not all cell mutations and fight against these mutated cells. For cancer to occur it is not only sufficient for our cells to irregularly proliferate but for all these mechanisms that attack the mutated cells to also be disrupted. As Sompayrac (2004, p. 7) emphasizes: "The current view is that to become a cancer cell, multiple growth-promoting systems must be activated inappropriately, and multiple safeguard system must be inactivated.". Thus, we can conclude that cancer does not arise independent of any other bodily regulatory mechanisms but is actually a glitch in those mechanisms designed to protect us. Analogously with the previous examples, cancer occurs when the organism cannot achieve homeostasis on its own. Millions of cancerous cells occur in our bodies but as long as our bodies neutralize these occurrences, we remain healthy. Once the mutated cancerous cells start to grow in number and out of normal pace and when our bodies' cancer-regulating systems malfunction we have a problem on our hands. This is the instance when the organism is harmed because it cannot bring itself to homeostasis, it cannot deal with the 'problem' of cancerous cells on its own. Here is when medical attention is sought. The treatment for cancer involves killing off as many cancerous cells that there are in order for the body to take hold of them and to repair on its own. In other words, medicine helps the body regain homeostasis. Doctors work together with the body, they aid nature in trying to treat cancer. However, cancer is a vicious disease and the treatments such as chemotherapy and radiation therapy have severe and sometimes lethal side effects. The important thing is that the underlying principle behind cancer follows the same rationale as previous examples. We are harmed when our organism cannot achieve or regain homeostasis itself. Medicine helps us by introducing resources that help the organism combat disease.

Here I have shown how the homeostatic view of medicine works to explain three conditions – leg fracture, bacterial infections and cancer. I will not go into

pathogenesis of other conditions but I think that the same holds for other conditions as well. If we take HIV the problem or harm in it is that our immune system is compromised so that we are prone to infections and diseases we otherwise would not have been affected by. The ability of our organism to achieve homeostasis is impaired. The rationale I have laid out earlier is behind most if not all pathological conditions. I suppose it works also for viral infections, autoimmune diseases and many other kinds of disorders.

In adopting the concept of homeostasis we can explain a wide range of medical conditions and disorders. While the idea of reductionism, breaking down and analyzing smaller and smaller components of an organism, is present and dominant in the contemporary medicine, homeostasis seems to be present and represented as well, although more implicitly than reductionism. As I have argued in the methodology section, the first line of business of philosophers is to make what is implicit explicit. Here it means bringing the ideas of homeostasis to the forefront and seeing how the idea is implicit in various fields of medicine and medical endeavours. I have also introduced concepts such as achieving and maintaining homeostasis. Let us not consider these as necessary and sufficient conditions of health in medicine but align to Strawson's ideas that these concepts and their mutual connection helps to illuminate the ideas behind health and disease in medicine. So far this has been the descriptive analysis meaning that I have tried to show how the idea of homeostasis is de facto implicit in medical reasoning. It is the first of our methodological steps of defining the explicandum, or outlining the concept of harm that we can start from in our analysis of harm in psychiatry. I do also believe that the idea of homeostasis should be in place. Homeostatic view of medicine makes sense. I will not go into the normative dimension here of what should be the idea of harm in medicine but I will turn to psychiatry and how can these considerations laid so far be implemented in our psychiatric thinking.

Now that we have made, following Strawson's methodological ideas, the theoretical rationale of medicine in terms of homeostasis explicit, our next order of business is applying this analogy the field of psychiatry and seeing how the concept of harm should be constructed. Here we are going from a descriptive project to a prescriptive, creative project of coming to harm in psychiatry.

### 6.4 Harm in Bodily Medicine vs. Harm in Psychiatry

In applying the analogy of harm as inability of the body to achieve and maintain homeostasis from bodily medicine to psychiatry, it is important to note important dissimilarities between bodily medicine and psychiatry. This can also be described as one of the biggest problems of psychiatry as a medical science, or rather its greatest challenge. And it has to do with psychiatric nosology.

A sobering moment in the recent history of psychiatry that revealed the psychiatric crisis happened in 1973 when an article published in Science described the experience of twelve subjects who pretended to 'hear voices' in order to try to be admitted to a psychiatric institution (Nesse & Stein, 2012, p. 1). They were hospitalized and received a diagnosis of schizophrenia regardless of the fact the they acted normal after admission (Nesse & Stein, 2012). As Nesse and Stein (2012) write, psychiatry aspired to gain recognition as a credible scientific field in the late 1970s, however, the lack of reliability in its diagnostics posed a significant obstacle. Then, new diagnostic manuals emerged such as DSM III and DSM IV which introduced "checklists of operationalized indicators" like those that are used to this day in psychiatric diagnosis (Nesse & Stein, 2012, p. 1). If we take any psychiatric condition there are checklists of symptoms for each condition a certain number of which have to be present in the patient for them to be diagnosed with a disorder. This is how both DSM and ICD are structured. There are many problems with this type of diagnosis. One of them is a high rate of comorbidity associated with psychiatric disorders meaning that "most individuals who have one disorder also qualify for additional

diagnoses" (Nesse & Stein, 2012, p. 2). This is diagnostically problematic because it might indicate that psychiatric categories are not precisely delineated. We might not 'cut nature at its joints' when it comes to psychiatric disorders and our categories might not turn out to be robust and valid as we might think they are. Another problem is the "heterogeneity of patients within diagnostic groups" (Nesse & Stein, 2012, p. 2) which means that in patients diagnosed with a particular condition the sets and combinations of symptoms that patients can have might be different to the point that two patients might not share any commonalities. We would generally expect in patients with the same diagnosis regularities and almost similar manifestation of symptoms. After all, that is how we can identify the disorder, predict its course and find a treatment. Once you take that away diagnosis starts to look more and more like a stab in the dark. Finally, the problem that is most revealing of the disparity between psychiatry and the rest of the bodily medicine is that "with the exception of neurological disorders such as Huntington's Disease, not one of the main DSM mental disorders can be validated by laboratory or imaging biomarkers." (Nesse & Stein, 2012, p. 2). While in the rest of medicine the most important and potent diagnostic tools are laboratory and imaging results, psychiatry is almost completely devoid of them. As Nesse and Stein (Nesse & Stein, 2012, p. 2) emphasize:

> Many medical disorders are defined by a specific etiology, or by distinctive anatomical or molecular abnormalities. Despite exhaustive searches, no comparable objective indicators have been found for any major mental disorder [27]. Some statistically significant neurobiological differences characterize certain diagnostic groups (for example, on brain imaging), but they are neither specific nor sensitive enough to validate any diagnosis. The chair of the DSM-IV Task Force, Allen Frances, notes "the

> disappointing fact is that not even one biological test is ready for inclusion in the criteria sets for DSM-V" [28].

Thus, there are considerable differences in diagnosis between psychiatry and the rest of medicine. I contend that the most important underlying reason is that there is still a lot that we do not know when it comes to psychiatric disorders. Our knowledge of mechanisms underlying psychiatric conditions, as well as pharmacotherapy in treating those conditions is unfortunately quite limited. As it is written in the DSM IV (American Psychiatric Association, 1994):

> In DSM-IV, there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder. There is also no assumption that all people described as having the same mental disorder are alike in all important ways.

Applying the holistic, homeostatic approach which I have showed is in place in bodily medicine, thus, might not work as well for psychiatry. Since we do not know enough about mechanisms in place in psychiatric disorders and how psychiatric disorders actually work, why they come about and how to treat them, it is hard to say when the organism is in the state of homeostasis and when it is not. Until we can identify the mechanism that fails, the way it fails and what can be done about it, applying the analogy from bodily medicine will not get us very far. I am by no means saying that it cannot be done. With diligent and systematic research that is underway in psychiatry we will surely know more and more about mental disorders, from biological, pharmacological, psychological and social perspective. I garner optimism and sincere hope with regards to such developments in psychiatry. However, until

these things take hold we have to be cognizant of the limitations of our current knowledge. As we have seen, harm in bodily medicine can be explicated almost completely in biological and scientific terms by evoking the holistic, homeostatic view. Harm can be explained using the concept of homeostasis, specifically it amounts to failure of a mechanism to achieve and maintain homeostasis using internal resources. Taking into considerations differences between psychiatry and bodily medicine, simply applying the homeostatic view to psychiatry does not seem tenable, at least not yet. Therefore, it is necessary to find alternative ways to conceptualize harm in psychiatry while still adopting and acknowledging some of the same principles that I have outlined are in place in bodily medicine. In the following chapter I lay out an account of harm in psychiatry. This is where the process of explication takes hold as we are going from a harm in bodily medicine as ability to achieve and maintain homeostasis to try to come up with its psychiatric counterpart. In doing so there are several principles that I have mentioned that should be adhered to. Following the principles of compassionate revisionism, the harm criterion should capture the patient's perspective as fully and faithfully possible while also being intelligible to the practices and modes of psychiatric practice. In addition to that, our concept should be constructed to follow the Carnap's guidelines of fruitfulness, exactness, similarity and simplicity. In the following chapter I construct the concept of harm that respects and follows these guidelines.

## 6.5 Conclusion

In this chapter I set the groundwork necessary for coming to the concept of harm in psychiatry. I lay out the methodology of such a project, deciding on the method of philosophical explication. The first step in the method is what Rudolf Carnap calls 'formulating the problem'. In this I help myself with Strawson's idea that the role of philosophy is to make implicit concepts explicit, and his idea of conceptual analysis as connectivism. I outline a concept of harm in medicine that relies on the homeostatic

ideas of health and propose a concept that consists in the inability of an organism to achieve and maintain homeostasis. I compare this view of harm to the needs and context of psychiatry. I conclude that this homeostatic view of health cannot be entirely transposed to harm in psychiatry. Thus, in the next chapter I construct the account of harm that would be suitable for psychiatry but that follows the same underlying principle as harm in bodily medicine.

# 7. An Account of Harm in Psychiatry

## 7.1 Introduction

In the previous chapter we have come to explicate what harm in medicine consists in. We have ended with the conclusion that harm in medicine, i.e. the homeostatic approach, cannot be quite transposed to psychiatry. Thus, we need an alternative way of formulating harm in psychiatry. In doing so it is important to stay close to the ideas of bodily medicine that I have laid out. If we are not able to carry the idea of harm over from bodily medicine to psychiatry, let us then try to find the next best, closest approximation to it.

In this chapter I argue that harm in psychiatry can be configured as 'inadequate resources in dealing with the problems in living'. While the diagnostic process in bodily medicine is very different from psychiatry there are useful psychological concepts that are used and that garner significant explanatory power. Lastly, I propose that this view of harm is compatible with a theory if well-being that consists of positive causal networks (PCN's) outlined by Michael Bishop.

## 7.2 Harm as inadequate resources in dealing with the problems of living

I argue that harm consists in *inadequate resources in dealing with the problems of living*. I will dedicate the remainder of this chapter to uncovering and explaining each facet of the claim and ultimately, how all the elements work together.

I will first deal with the concept of 'resources'. Resources can be defined as "those entities that either are centrally valued in their own right (e.g., self-esteem,

close attachments, health, and inner peace) or act as a means to obtain centrally valued ends (e.g., money, social support, and credit)." (Hobfoll, 2002, p. 307). Resources can be further classified as "distal and proximal to the self, internal and external, and biological and cultural" (Hobfoll, 2002, p. 307). Internal resources are those such as self-efficacy, dispositional optimism, high self-esteem, resilience, personality hardiness, emotional regulation, interpersonal and communicative skills, problem-solving skills and so on. This is a large pool that cuts across many psychological views and theories of resources. For the purposes of this work, the take-home is that they are internal to the person and these are the ones we are interested in. These are traits, dispositions, moods, emotions, attitudes, ways of thinking and feeling. Conversely, external resources are things like money, social status, social support and the environment that the person lives in. While these external factors are important and are integral to the well-being of a person, we are not looking at them when diagnosing patients. Diagnosis does not rely on person's external resources. For diagnosing persons it is not important how much money the person has, but whether the person has internal resources conducive to making a living for themselves, helping them become accepted, functioning and valued member of the society. The person that suffers from a condition such as schizophrenia probably has a low-earning potential but that is not the problem we are trying to remedy with psychiatric intervention. What we are trying to do is help the person obtain resources for a healthy, functioning and fulfilling life which would foster further acquisition of variety of resources. It is my assumption that in developing internal resources one also develops the abilities to acquire external resources. For example, it seems commonsensical that people who have high self-esteem more easily approach others, make friends and gain social support than people of low self-esteem.

Thus, in the process of diagnosis we are dealing with the particular condition that a person has and with respect to harm determining whether the condition is

harmful in itself we are looking at internal resources. We are determining whether the condition is in virtue of its characteristics harmful. Recall that when we talked about depathologization of homosexuality we said that it is a condition that is not harmful in itself meaning that same-gender sexuality, love and relationships are not internally problematic because many can live happy and fulfilling lives. The aspect that is harmful is the disapproving and stigmatizing social response and climate to homosexuality. Thus, homosexuality is not internally harmful but harm comes externally from social responses.

One the main goals of using the concept of 'resources' is to interface certain psychological with selected psychiatric notions which I will elaborate on in the following. In using the notion of resources, we can connect the psychiatry with the vast psychological literature that studies human mind and cognition. Let us see how that can be done.

The idea of resources is by no means novel in psychology as there is a vast body of scientific studies and papers that has accumulated over the years. Hobfoll (2002) in one of his articles lays out what is common to all the different ideas and theories of resources that have been postulated thus far. I take his analysis as indicative of what we should think and how to conceptualize the notion of resources. The first of these common postulates is that "people strive to obtain, retain, protect, and foster resources in biological, cognitive, and social domains." (Hobfoll, 2002, p. 317). This means that in most of these theories on resources, they are considered to be something valuable. As such people strive to obtain them, keep them, protect and maintain them. Not only that but people seem to act as to minimize any future resource loss by accumulation of resources (Hobfoll, 2002, p. 317). The second thing that seems to be common is that "people with resources are less likely to encounter stressful circumstances that negatively affect psychological and physical well-being (King et al., 1999)."(Hobfoll, 2002, pp. 317–318). It is postulated that those who obtain

and are able to maintain resources are steered by their resources away from situations that might compromise their well-being. Think for example students that have acquired discipline and work ethic. They have developed a habit of doing things in a timely manner and as a result they minimize the occurrence of stressful situations of doing things at the last minute, not being secure in their work because of lack of preparation or lack of research that they might have needed to do. Another 'mechanism' of resources

> is that those who possess resources are more capable of solving the problems inherent in stressful circumstances. In this regard, stressful circumstances can be seen as a life puzzle that must be unraveled. If they have more resources, people are more likely to have either the specific resources needed to fit demands or resources that can provide them access to the resources that fit demands. (Hobfoll, 2002, p. 318)

Resources do not fall out of the sky but they have to be acquired at one point or another in one's life. A way to obtain resources is to develop them through dealing with challenging situations. For example, a way to gain resilience is to persist and persevere in difficult situations, to endure hardship, pain or discomfort in reaching some goal. After such experiences the person's outlook on stressful situations can become that of a challenge rather than a problem. The person can deal with stressful situations easier and have more knowledge and understanding of how to surpass them. Furthermore, they may acquire other resources more easily since they have gotten used to dealing with difficult and tedious processes of acquiring resources.

Another regularity that is present across board when it comes to resources is that "resources are linked to other resources. Hence, there is a general tendency for enrichment of resources among those who possess a solid resource reservoir."

(Hobfoll, 2002, p. 318). Resources facilitate future proliferation of resources. As I have indicated in the example above on resilience – once you acquire resilience it is much easier for you to acquire any other such resource.

Furthermore, and linked to the previous points, "the influence of resources is long term and tends not to be transient as with the impact of stress." (Hobfoll, 2002, p. 319). In explaining this, Hobfoll introduces the idea of resource caravans which means that "resources, or their lack, tend not to exist in isolation, but rather will aggregate such that, for example, individuals with high self-esteem will also possess a stronger sense of mastery and have better functioning social support systems" (Cozzarelli, 1993; Rini et al., 1999) (Hobfoll, 2002, p. 312). Similar to some of the previous ideas on resources, it seems that resources in general facilitate the accumulation and maintenance of other resources. Then, these resources are applied in various situations and utilized in different ways to manage challenging circumstances. This is possible especially because "resources come in bundles" and "their impacts tends to hold across time and different circumstances" (Hobfoll, 2002, p. 319). Thus, the person that has a significant amount of resources not only has the upper hand at acquiring new resources but they also have a selection of resources to apply to all kinds of life challenges.

Additionally, "resources become valued in their own right, and those who possess resources (e.g., support from intimates and money) are viewed by others, and will view themselves, more favorably." (Hobfoll, 2002, p. 319). So far the possession of resources was described as having an instrumental value in people's lives meaning that they are make dealing with live situations and problems easier and more efficient. But resources also, and in part, tend to be valued intrinsically. Hobfoll mentions that resources are immersed and valued in part with respect to the person's culture. We can recognize resources that are socially and culturally valued. For example, Eastern European cultures put considerable value on family relations,

loyalty and caretaking of family members. In Northern European countries individualism, independence and self-sufficiency are characteristics that are emphasized and valued. Thus, Hobfoll argues that resources in these cultures will be sought both because they alleviate and help with stressful situations, and will be sought in their own right, seen as something valuable in itself. The possession of these resources will positively affect the person's social standing in a particular culture and their view of themselves. Regardless of people's cultures, we can talk of the similar effect in individualistic terms – in terms of personal values. If a person values hard work and resilience in the face of life's hardships, in dealing with life in such manner it will add to the person's positive image of themselves. Even in the culture that maybe puts emphasis on other aspects of the human condition, the fact that the person themselves values these things and works towards them will undoubtedly positively affect their life.

Let us then sum up the generalities concerning resources. Resources seem to be something that is sought for as people strive to acquire and maintain them. People who possess resources tend to be steered away from difficult and stressful situations because of them. When confronted with difficult and stressful situations, resources are the things that is relied on and which aid people in overcoming such situations. Resources tend to facilitate the acquisition of other and new resources and they tend to exist in resource bundles or 'caravans'. These congregation of resources help in solving problems of various nature and difficulty meaning that their application is transient and not necessarily domain related. Finally, attainment of resources is in part intrinsically valued, in their own right, not just as means of dealing with stress but as an end in itself. As something worthy of attainment and preservation for its own sake and which positively affects the person's social standing and their self-image.

Now that we have an idea of resources represent. Let us look at what 'inadequate resources' amount to in the definition of harm as inadequate resources in dealing with the problems of living. Inadequate in my account can mean a couple of things. One interpretation of inadequate is insufficient. This means that the resources that the person has are insufficient with regards to the gravity, intensity, and seriousness of their problems. This can come about in one of two ways. It either means that the person has not acquired sufficient resources to deal with life problems. Take for example persons who have been raised in a sheltered, protective and co-dependent family environments. As a result, in coming of age they might have hard time adopting to adult life and have problems gaining independence from their parents. In other words, they have not evolved resources that help them deal in a constructive and independent ways with the world. Adult life may seem as a hardship for them, they may feel lost, and inept at dealing with situations that require from them responsibility, independence and autonomy. They might be introduced into the psychiatric and/or psychotherapeutic process to help them develop the resources for coping with life's challenges which can be done through psychotherapy.

Another reason why a person might not have enough resources to deal with the problems of living are tragic changes in intensity, severity or longevity of their problems of living. When the problems of living take an unexpected and dramatic turn, people might not have the abilities to cope with them. Being diagnosed with a terminal illness, loss of a family member, war, hunger and incarceration are just some examples of these extreme changes in problems of living that can have a striking effect on the person's psyche and can trigger psychiatric conditions.

This is especially poignant in trauma survivors and one of the virtues of this theory of harm is that it can account for their experiences. Traumatic experience such as rape, mugging, witnessing a murder and so on are situations where we see a steep uprise in severity, and complexity of problems of living. However, the person may not

yet have the resources that would mitigate the effect of these problems. Trauma survivors are thus introduced in the psychiatric context to help them get to the point where they can deal and process what has happened to them. In this they may be aided by pharmacological and/or psychotherapeutic treatments.

Another way in which resources might be inadequate is that they may be the 'wrong' resources for dealing with the problems of living, i.e. they may be maladaptive. To get a clue about what maladaptive resources amount to let us take a look at the definition of maladaptive behaviors. Maladaptive behaviors are defined as "actions or tendencies that don't allow an individual to adjust well to certain situations." and "they are typically disruptive and dysfunctional behaviors which can range from mild to severe in scope." (*Maladaptive Behaviors Definition*, n.d.). The idea behind maladaptive behaviors is that they are adopted to deal or cope with a certain situation or a stressor but are bad for the person and may make a situation worse. For example, alcoholism and substance abuse in general might be introduced to cope with feelings of depression, social anxiety or boredom but they make the situation worse without solving the initial problem or acquiring an adaptive response to it. As it is stated "Maladaptive behaviors are typically used as a means of reducing mental discomfort and anxiety but are not effective in this regard and can sometimes even make it worse." (*Maladaptive Behaviors Definition*, n.d.). Here we are talking about maladaptive behavior which may be just one psychological facet while resources refer to a whole set of attitudes, behaviors and mechanisms that people employ in dealing with problems of living. Thus, maladaptive resources are those that people use to deal with stressful situations but are actually not effective in dealing with such situations and might even bring more harm. Alcohol and substance abuse is just one example. We can also consider the case of psychopathy or antisocial personality disorder where it seems that people characterized by such conditions seem to 'get in trouble', put themselves in situations that are dangerous for themselves and others. Their

resources might be violence, aggression, manipulation, lying, cheating or some other behaviors that are corrosive to their own well-being and to their relationships with others.

Now that we have a better understanding of both what 'inadequate' and 'resources' signify let us turn our attention to what 'dealing with' means in our theory of harm[20]. 'Dealing with' is basically the utilization of resources in dealing with the problems of living. It is using the resources one has at their disposal to solve some problem or a stressful situation. However, this is also where things can go wrong where it can be the case that the person actually has the right resources but has trouble utilizing them in. This can happen because of the problems in processing, cognitive overload and so on. One such example can be found in autism where it is postulated that

> the combination of human stimuli and the enforced global processing in this condition, resulted, according to the complexity theory, in a processing demand overload, reflected by the difficulties individuals with ASD showed in processing context. (Ben-Yosef et al., 2017, p. 526)

The problem with autism seems to be that they are under significant cognitive processing overload which disables them from utilizing the resources efficiently even if they have them. The theory states the problem with autism is that most of processing that people who do not have autism go through is too cognitively taxing with persons with autism. As a result, in people with autism their processing mechanisms are overwhelmed and cannot be properly utilized to deal with everyday challenges.

[20] I have defined harm earlier as inadequate resources in dealing with the problems of living.

Finally, the last piece of the puzzle when it comes to the idea of harm as inadequate resources in dealing with the problems of living, is the 'problems of living'. This concept has been present in the psychiatric literature in Thomas Szasz (1960). Recall that Szasz argues that mental disorders are a myth. What we consider as mental disorders is either a dysfunction in the brain, a brain lesion which falls under the domain of neuroscience and neurology, or it is a problem of living which does not fall under the domain of medicine at all. Similar to Szasz, I do think that mental disorders reflect inabilities in dealing with problems of living. However, I think these rightfully belong to the field of medicine and psychiatry.

A problem can be defined as an obstacle in one's life that takes some kind of effort in overcoming it. Thus, a problem can be getting up in the morning, finding a job as well as dealing with cancer. The difference between these is rather obvious. What sets them apart is that they differ in intensity, longevity, difficulty and other numerous characteristics. However, what they have in common is that they are problems. They represent a kind of obstacle or challenge in daily lives of people dealing with them. Granted, the resources and efforts employed in solving them will naturally also differ. Yet, they are all problems nonetheless.

What is important and elegant about the concept of 'problems of living' is that it is a universally human category, a human condition shared by all humans alike. All people have problems, regardless of their age, gender, sex, geographical location, race, occupation and so on. All exert effort in solving them, being cognitive, physical, emotional and so on. All rejoice in overcoming them. And maybe even learn and grow in the process.

As long as the resources the person has outweigh the gravity of their problems, and they are able to mitigate the negative effects of problems, the person is not harmed. However, when the problems of living outweigh the resources the person has, the person becomes harmed.

Life expects from us constant negotiation between resources that we have and the problems of living that we are dealing. Sometimes we lack resources but are able to build them up fast according to situations and based on other resources which we already possessed. Sometimes the problems of living never get challenging to the point that the person has to acquire an extensive amount of resources in dealing with them. All lives are different, but what might be similar is that all utilize some resources in dealing with the problems of living.

I contend that this account of harm satisfied much of desiderata we have set out to include in our analysis of harm. Following the principles of revisionist compassionatism I believe the formulation of harm as inadequate resources of dealing with the problems of living is designed to capture the patient's perspective as fully and faithfully as possible. In order to identify whether the person is harmed we need to gain an understanding of the resources the person has as well as the problems of living they are dealing with. This undoubtedly requires extensive interviewing and real, genuine understanding of the person's position. The understanding of elements in such an account of harm with respect to individual's life probably cannot be achieved without thorough examination of the patient's perspective.

Furthermore, even though the account aims to fully capture the patient's perspective, it aligns to the precepts and ideas of psychiatric practice as well. Interfacing the concepts used – inadequate resources, dealing with the problems of living, with psychological concepts, studies and data offers theoretical and scientific backdrop against which to think, judge and assess patient's perspectives. It helps to determine the way and extent to which the person is harmed. It offers a theoretical framework that is helpful and perspicacious in determining the patient's state.

When it comes to the method of explication I believe this account of harm satisfies all of the requirements – fruitfulness, exactness, simplicity and similarity, proposed by Rudolf Carnap (1962). Recall that in the previous chapter I have outlined

harm in bodily medicine which consists in the inability of an organism to achieve and maintain homeostasis. Explained generally and in lay terms, harm occurs when the organism cannot 'deal with the problem on its own' and thus may need medical assistance in order to restore homeostasis. The account of harm that I have proposed in psychiatry as 'inadequate resources in dealing with the problems of living' follows the same underlying principle as that of harm in bodily medicine. When we cannot 'deal with problems on our own', meaning by utilizing internal resources that we have at our disposal, then we are harmed. Thus, I believe the criteria of similarity to the explicandum of harm in bodily medicine is satisfied.

Next, let us see whether this concept of harm is fruitful. Fruitfulness for Carnap involves "allowing the formulation of empirical generalizations or theorems in the case of logic and mathematics." (Biturajac & Jurjako, 2022, p. 242). However, in recent times, the domain of fruitfulness has broadened enabling us to configure concepts which would satisfy "practical and ethical desiderata of our intellectual inquiries" (Biturajac & Jurjako, 2022, p. 242)(see also Dutilh Novaes (2020)). Fruitfulness for Carnap is arguably the most important criterion. And, fruitfulness is judged with respect to the area of inquiry we want the target concept, explicatum, to operate in. I believe the concept of harm as inadequate resources in dealing with the problems of living is fruitful in that it satisfies our criteria of what compassionist concepts should look like. The role of the concept of harm is to represent the patient's perspective while also keeping to the expertise of the psychiatric practice. The concept captures the patient's perspective as it is impossible to judge whether a person is harmed without hearing their own experiences and stories from their own point of view. However, the account of harm is also entrenched in the rules and ideas of psychiatric and psychotherapeutic practice as it is interfaced with psychological concept such as resources. I believe it requires the practical requirements of psychiatric practice as it points to things to look for in interviewing and assessing states of patients. It may

even be operationalized as the subject of internal resources has been thoroughly studied and examined by psychology. Judged by the standards of explicandum which we have found in bodily medicine, as I have explained, this version of the concept of harm is much more suitable and fruitful for the psychiatric context than the inability to achieve and maintain homeostasis, although it is based on the similar underlying principle.

When it comes to exactness, the harm as inadequate resources in dealing with the problems of living might not seem all that exact. However, we have to be cognizant of the level of generality we are dealing with, it being the formulation of harm that would be applicable to most, if not all psychiatric conditions. It is expected that this account would be broad and all-encompassing. However, this concept of harm can be applied and translated into specificities. It affords itself to explanations of specific conditions even though it is a general concept as I have shown earlier in this chapter. I believe it suits the needs and interests of exactness at the level which it is supposed to operate on.

When it comes to simplicity, the concept of harm as inadequate resources of dealing with the problems of living. With respect to the explicandum from bodily medicine I would say that it is neither more nor less simple than it. I do think it is simple enough and if I daresay an elegant solution that nicely cuts across psychiatric conditions while keeping to the precepts of psychiatric practice.

### 7.3 Harm and Well-being

In the discussion on mental disorders, harm is sometimes explained or described in terms of well-being. Harm is then configured as a lack of well-being or some kind of deprivation. In chapter on eliminativism I have shown that the concept of well-being cannot be entirely equated with harm by showing that benefit and harm are not

completely antonymous (see chapter on Eliminativism). However, to dismiss the ideas of well-being entirely in talking about harm might seem a bit premature.

Several authors have made connections between harm and well-being, even defining harm in terms of the theories of well-being. Muckler and Taylor's (2020) paper attest to this as they analyze harm by taking Parfit's (1984) threefold grouping of theories of well-being - objective lists theories, desire-satisfaction theories and hedonistic theories. Objective list theories propose that there are some objective things that make life worth living and well-being consists in attaining and nurturing them, for example, family, meaningful work, friends, hobbies and so on. Desire satisfaction theories state that well-being basically consists in getting what you want – having your desires fulfilled. Hedonism is a theory according to which what is non-instrumentally good for an individual is the obtainment of pleasure.

Muckler and Taylor (2020) take particular conditions and test them through each of these theories to arrive at the conclusion about whether the condition is harmful. A condition is harmful if according to most (if not all) of these theories of well-being it results in harm. In lack of other substantive and elaborate theories of harm in psychiatry this seems as a sensible philosophical move. However, there are two problems with it from the standpoint of harm analysis. First, Parfit's distinction between these three theories – objective lists, desire-satisfaction and hedonism while presenting the foundation of the contemporary discussion on well-being, is somewhat outdated. There are many more theories and taxonomies of theories that have been developed since Parfit's *Reasons and Persons* (1984). Not only that but many different iterations and adjustments to the theories have been made in light of various counterexamples. Guy Fletcher provides a nice overview of them in his *The Philosophy of Well-Being* (2016) but even that can be considered a bit outdated as the discussion on well-being is lively, proliferous and very much alive still. Applying the Parfit's threefold distinction would be akin to talking about knowledge in terms on

the tripartite theory of knowledge as justified, true, belief. It seems to me that the philosophical state of the art has surpassed and outgrown that. The second reason why using the Parfit's threefold distinction is not the best idea is that it is philosophically inelegant. Anyone could object to them by asking why these theories in particular would assume the importance in thinking about harm. Defining harm in terms of these three is theory-laden and many would disagree or object to at least one, if not all three of these theories of well-being which is how the discussion on well-being developed into the vast and prominent discussion that it is today. The contemporary trend in thinking about well-being seems to be finding common denominators between most of the theories of well-being while also taking into account vast body of scientific evidence on well-being from psychology, economics and related scientific fields (Alexandrova, 2017; Bishop, 2015; Fletcher, 2016; Tiberius & Plakias, 2010). In thinking about harm in psychiatry, I do not want to resort to defining it in terms of a theory of well-being as I think harm can be a standalone concept. However, it is possible and prudent for it to be compatible with our ideas on well-being.

There is an account of well-being that fits elegantly with the account of harm I have proposed. It is a procedural account that tries finding common denominators in theories of well-being and is consistent and informed by scientific inquiries on harm. Rather than treating well-being as a separate and disparate way of accounting for harm I argue that we can take harm and well-being in unison by devising a whole conceptual system that is coherent and consistent as well as informative. To do this we need a theory of well-being that is empirically adequate as well meaning that it seriously takes into account the precepts on which I base my account of harm, namely in terms of psychological concepts and findings, such as those on resources I have outlined earlier. Here is where the network theory of well-being comes in.

The network theory of well-being states that the "state of well-being is the state of being in (or to use philosophical jargon *instantiating*) a positive causal network" (Bishop, 2015, p. 10). This means that a "person high in well-being has positive emotions, attitudes, traits, and accomplishments that form an interlocking web of states that build and feed on each other." (Bishop, 2015, p. 10). To have well-being is being in this feedback loop of positive emotions, attitudes, traits, and accomplishments.

Let us look at an example of Well-being Wellington. Wellington works at a job he likes and is good at, teaching high-school chemistry. He adores chemistry and he even finished a PhD program in chemistry. However, he always wanted to be a teacher and that is how he feels he is making a difference. Besides chemistry being fun for him Wellington has a sense of mastery, and he exerts self-esteem when it comes to teaching it. Students recognize that and respect his knowledge and enthusiasm. As he goes to the job he likes and excels at his work, he is an esteemed colleague and enjoys respect both from students and colleagues. Since going to work is not a hassle but a pleasure, he is often in high spirits smiling, joking with his colleagues and he is ready to jump in when some of his colleagues cannot make it to work. As a result, he is well-liked and has a couple of close friends among his colleagues. Some of these friends enjoy hiking and spending time in nature over which they bonded with Wellington in the first place. On weekends they go on fieldtrips with mountaineering and hiking clubs. Wellington gained more valuable friends there and that is also where he met his boyfriend Mark. Mark is a botanist and Wellington was taken aback by his knowledge of the natural world which always fascinated Wellington. On one of these fieldtrips Wellington came across a stray dog who was in bad shape. He decided to take care of her which resulted in him adopting her. Her name is Maggie. Now Wellington, Mark and Maggie go for long walks in nature.

We can observe in the case of Wellington that one good thing led to another. His positive emotions, traits, attitudes and accomplishments facilitated and enforced their further proliferation. To the untrained eye Wellington might just be blessed by a serious of fortunate events but under closer inspection we can see that is not the case. Wellington's state of well-being has a lot to do with his internal characteristics, his decisions and choices and the external environment he puts himself in. His positive attitude towards his work facilitates connection with his co-workers and impacts his work-satisfaction. This also translates to his personal life which also flourishes as a result of his positive attitudes, emotions, traits and accomplishments. Wellington's case just illustrates this positive feedback loop that is characteristic of positive causal networks.

Since I have defined harm in psychiatry as inadequate resources in dealing with the problems of living, let us see how this relates to the network theory of well-being proposed by Bishop (2015). We can see that many of the properties of resources as defined in psychology resemble and can be integrated in the network theory of well-being. Resources essentially are traits, emotions and attitudes. And as we have seen the network theory of well-being is a feedback loop of positive traits, emotions, attitudes and accomplishments. It is important to emphasize that these resources in terms of harm in psychiatry have to be internal since diagnosis of patients is done with regards to their internal states and conditions. Thus, lack of money cannot be considered a lack of resources in psychiatric terms as lack of money represents the lack of some external value. However, lack of money can be a result of some internal condition that interferes with people's livelihood for example, gambling addiction. Thus, we can reconfigure the network theory of well-being as a feedback loop of 'positive' resources, i.e. of internal resources that aid people in dealing and exceling in problems of living.

Now, it could be argued that there is a disparity between harm and well-being. Having adequate resources to deal with the problems of living might not necessarily imply that one is in a state of well-being. We can imagine a person that has resources to deal with the problems of living but is not living life filled with well-being. As Bishop says, they can be in some kind of neutral causal network of traits, emotions and attitudes. Thus, the idea of reducing harm by gaining adequate resources in dealing with the problems of living does not necessarily bring one into a state of well-being.

However, there are some characteristics postulated that indicate that gaining adequate resources might significantly increase likelihood of that person having well-being. For example, solving a problem tends to increase one's resources with dealing with problems as well as give one a sense of accomplishment and raise their self-esteem. This tends to be motivating and translates to other domains of life where there is positive feedback with regards to the person's newfound level of self-esteem and sense of accomplishment. Then, once the person sees positive effects and gains positive feedback from the environment the person approaches challenges and problems more confidently and with more positive emotions. In approaching problems from that attitude, the person has more 'tools' at their disposal, as positive emotions tend to "broaden people's thought-action repertoires, encouraging them to discover novel lines of thought or action." (Fredrickson & Joiner, 2002, p. 172). In doing this, the person gains more resources to have at their disposal as well as gains positive emotion from utilizing them. And the process goes on and on. This is what some authors refer to as an 'upward spiral' (Bishop, 2015). Barbara Fredrickson proposes a Broaden and Build hypothesis that is a strong case for the positive causal network theory of well-being and for the phenomenon of an 'upwards spiral' (Bishop, 2015). It is evocative of the idea that positive emotion and resources attainment correlates with further positive emotion and resources attainment, which in turn correlates with one's well-being: "As individuals discover new ideas and actions, they

build their physical, intellectual, social, and psychological resources." (Fredrickson & Joiner, 2002, p. 172) meaning that their resources repertoire deepens and widens. As a result of that, people have more resources in approaching problems of living which makes them deal with them more efficiently, with less effort and more prowess. It creates a feedback loop of acquiring resources and positive emotion which perpetuates itself.

Let us imagine a person called Abe who wants to quit drinking. Going through one weekend without drinking gives him at least a tiny sense of accomplishment and optimism that he can gain control of his bad habit. However, now Abe has a bit of free time on his hands since he is not going on benders and curing his hangovers throughout the weekends. In order to stave off crises and keep himself sober, Abe decides to keep himself busy. He goes on long walks all across his city. Physical activity improves his mood, he sleeps better and he feels altogether better and healthier. He makes going on long walks a habit seeing the positive changes it brings him. Whenever he feels like drinking he goes for a walk and by the time he gets home he feels better and more optimistic about his decision of quitting drinking. Being sober for a while gives him more energy. Besides walking, he keeps himself busy by fixing things around his house and cleaning the garage. Now his place looks nicer and he is more comfortable, feels better living there. Looking around his place and seeing the positive changes he feels a sense of accomplishment and feels like his life has taken a much better, healthier turn. He does not go out as much with his drinking buddies as he used to since he fears he might not be able to control himself. He stays in on Fridays and Saturdays. He remembers a time when reading was his favorite hobby and so he reads in the evenings to keep himself busy before he goes to sleep. On one of his walks, he stumbles upon a pamphlet for a book club in his local library. He decides to join as the meetings are on Fridays which might help him stay sober and away from the bars. He meets people with similar interests that he starts hanging out with. They

sit in the library café, drink tea and discuss books on weekends. And life gets better and better.

The example of Abe, of course, just serves to illustrate the mechanism of 'upward spirals' and positive causal networks. In people who have problems with alcohol and alcohol addictions getting and staying sober is often a much more grueling endeavor. Their environment often enables and supports their maladaptive behavior, and it is often a vicious circle that is hard to get out of.

Thus, while the obtainment of adequate resources in dealing with the problems of living do not necessarily increase or ensures one's well-being, it highly increases one's likelihood of getting into a state of well-being and staying in that state.

## 7.4 Conclusion

This chapter is dedicated to the account of harm in psychiatry. I formulate harm as inadequate resources in dealing with the problems of living and explain each of these facets as well as how they work together in an account of harm. This is the creative and constructive part of philosophical explication. I analyse how this account of harm measures up to principles of revisionist compassionatism as well as to the precepts of philosophical explication expounded by Rudolf Carnap. I end the chapter in arguing that the harm proposed as inadequate resources in dealing with the problems of living is compatible with an account of well-being proposed by Michael Bishop that defines well-being as a positive causal network.

# Conclusion

Harm entered psychiatry in 1970s in the wake of depathologising homosexuality. This was an unprecedented move as it introduced and seriously taken into account the patient's perspective in psychiatry. Even though harm has been featured in the debate on mental disorders in philosophy of psychiatry, as well as in psychiatric manuals such as DSM and ICD, a serious and thorough examination of harm was left to be desired. This thesis remedies this by offering and extensive and comprehensive examination of the concept of harm in psychiatry. It keeps to wants, needs and ideas of contemporary philosophical discussions while working inside the context of psychiatry. Harm was and remains an important element of psychiatry. It captures the first person's perspective of the individual in the process of psychiatric care. It is the embodiment of the second of the three pillars of good psychiatric practice – scientific research, patient's perspective and psychiatric/clinical expertise. Thoughts, research and discussion on harm should be kept alive as it prevents some of the dangerous of psychiatric misuse and gives voice to the receivers of health care. The experiences, wishes, problems and ways of dealing with them of psychiatric patients can be understood and formulated with the help of the harm criterion I propose. Harm as 'inadequate resources in dealing with the problems of living' necessitates being receptive and cognizant of the patient's point of view. I believe it also stays true to the concepts and desiderata of the psychiatric practice and is commensurable and understandable to psychiatric expertise and is consistent with scientific research. The importance of harm should not be understated. Introducing harm in psychiatry has

been a revolutionary move and we should strive to further revolutionize the concept of harm with the changing times and advancements in psychiatric thought and practice.

# Bibliography

Agich, G. J. (1994). Evaluative judgment and personality disorder. In J. Z. Sadler, O. P. Wiggings, & M. A. Schwartz (Eds.), *Philosophical perspectives on psychiatric diagnostic classification* (pp. 233–245). Johns Hopkins University Press.

Ahmed, P. (2012). *State Mental Hospitals: What Happens When They Close*. Springer Science & Business Media.

Alexandrova, A. (2017). *A Philosophy for the Science of Well-being*. Oxford University Press.

Allport, G. W. (1954). *The nature of prejudice* (pp. xviii, 537). Addison-Wesley.

Althof, S. E., McMahon, C. G., & Rowland, D. L. (2022). Advances and Missteps in Diagnosing Premature Ejaculation: Analysis and Future Directions. *The Journal of Sexual Medicine*, *19*(1), 64–73. https://doi.org/10.1016/j.jsxm.2021.10.010

American Psychiatric Association. (1980). *Diagnostic and statistical manual of mental disorders: DSM-III* (2nd Printing edition). American Psychiatric Association.

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders: DSM-5* (5th ed.). American Psychiatric Association.

Amoretti, M. C., & Lalumera, E. (2019). Harm should not be a necessary criterion for mental disorder: Some reflections on the DSM-5 definition of mental disorder. *Theoretical Medicine and Bioethics*, *40*(4), 321–337. https://doi.org/10.1007/s11017-019-09499-4

Amoretti, M. C., & Lalumera, E. (2022). Wherein is the concept of disease normative? From weak normativity to value-conscious naturalism. *Medicine, Health Care, and Philosophy*, *25*, 47–60. https://doi.org/10.1007/s11019-021-10048-x

Ananth, M. (2008). *In defense of an evolutionary concept of health: Nature, norms, and human biology*. Ashgate. https://doi.org/10.4324/9781351155847

Baker, A. (2004). *Simplicity*. https://stanford.library.sydney.edu.au/archives/spr2017/entries/simplicity/

Bayer, R. (1987). *Homosexuality and American psychiatry: The politics of diagnosis*. Princeton University Press.

Ben-Yosef, D., Anaki, D., & Golan, O. (2017). Context processing in adolescents with autism spectrum disorder: How complex could it be? *Autism Research*, *10*(3), 520–530. https://doi.org/10.1002/aur.1676

Bhugra. (1989). Attitudes towards mental illness. *Acta Psychiatrica Scandinavica*, *80*(1), 1–12. https://doi.org/10.1111/j.1600-0447.1989.tb01293.x

Bigelow, J., & Pargetter, R. (1987). Functions. *Journal of Philosophy*, *84*(4), 181–196. https://doi.org/10.2307/2027157

Billman, G. E. (2010). The grand challenge of physiology: To integrate function from molecules to man. *Frontiers in Physiology*, *1*. https://www.frontiersin.org/articles/10.3389/fphys.2010.00005

Billman, G. E. (2013). Homeostasis: The Dynamic Self-Regulatory Process that Maintains Health and Buffers Against Disease. *Handbook of Systems and Complexity in Health*, 159–170. https://doi.org/10.1007/978-1-4614-4998-0_10

Billman, G. E. (2020). Homeostasis: The Underappreciated and Far Too Often Ignored Central Organizing Principle of Physiology. *Frontiers in Physiology*, *11*, 200. https://doi.org/10.3389/fphys.2020.00200

Bishop, M. A. (2015). *The Good Life: Unifying the Philosophy and Psychology of Well-being*. Oxford University Press.

Biturajac, M., & Jurjako, M. (2022). Reconsidering harm in psychiatric manuals within an explicationist framework. *Medicine, Health Care and Philosophy*, *25*, 239–249. https://doi.org/10.1007/s11019-021-10064-x

Blackburn, S. (2008). The Oxford Dictionary of Philosophy. In *The Oxford Dictionary of Philosophy*. Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780199541430.001.0001/acref-9780199541430

Bolton, D. (2008). *What is mental disorder? An essay in philosophy, science, and values*. Oxford University Press.

Boorse, C. (1975). On the distinction between disease and illness. *Philosophy and Public Affairs*, *5*(1), 49–68.

Boorse, C. (1976). What a theory of mental health should. *Journal for the Theory of Social Behaviour*, *6*(1), 61–84.

Boorse, C. (1977). Health as a theoretical concept. *Philosophy of Science*, *44*(4), 542–573. https://doi.org/10.1086/288768

Boorse, C. (1997). A rebuttal on health. In J. M. Humber & R. F. Almeder (Eds.), *What Is Disease?* (pp. 1–134). Humana Press. https://doi.org/10.1007/978-1-59259-451-1_1

Boorse, C. (2011). Concepts of Health and Disease. In F. Gifford (Ed.), *Philosophy of Medicine* (Vol. 16, pp. 13–64). North-Holland. https://doi.org/10.1016/B978-0-444-51787-6.50002-7

Boorse, C. (2014). A second rebuttal on health. *The Journal of Medicine and Philosophy*, *39*(6), 683–724. https://doi.org/10.1093/jmp/jhu035

Bortolotti, L. (2020). Doctors without 'disorders.' *Aristotelian Society Supplementary Volume*, *94*(1), 163–184. https://doi.org/10.1093/arisup/akaa006

Brody, H., & Miller, F. G. (1998). The Internal Morality of Medicine: Explication and Application to Managed Care. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *23*(4), 384–410. https://doi.org/10.1076/jmep.23.4.384.2572

*Broken leg*. (2017, October 18). Nhs.Uk. https://www.nhs.uk/conditions/broken-leg/

Brun, G. (2016). Explication as a method of conceptual re-engineering. *Erkenntnis*, *81*(6), 1211–1241. https://doi.org/10.1007/s10670-015-9791-5

Callahan, D. (1996). The Goals of Medicine: Setting New Priorities. *Hastings Center Report*, *26*(6).

Çaparlar, C. Ö., & Dönmez, A. (2016). What is Scientific Research and How Can it be Done? *Turkish Journal of Anaesthesiology and Reanimation*, *44*(4), 212–218. https://doi.org/10.5152/TJAR.2016.34711

Cappelen, H. (2018). *Introduction to conceptual engineering*. Oxford University Press.

Carnap, R. (1962). *Logical Foundations of Probability* (2nd ed.). Chicago Univ. Press. https://philpapers.org/rec/CARLFO-6

Carnap, R. (1971). *Logical foundations of probability* (4. impr). University of Chicago Press.

Cartwright, S. A. (1851). Report on the disease and physical peculiarities of the Negro race. *The New Orleans Medical and Surgical Journal*, 89–92.

CDC. (2022, January 3). *What You Should Know About Antibiotics*. Centers for Disease Control and Prevention. https://www.cdc.gov/antibiotic-use/q-a.html

Chin-Yee, B., & Upshur, R. E. G. (2017). Re-evaluating concepts of biological function in clinical medicine: Towards a new naturalistic theory of disease. *Theoretical Medicine and Bioethics*, *38*(4), 245–264. https://doi.org/10.1007/s11017-017-9410-3

Christensen, R. (2011). Involuntary Psychiatric Hospitalization and Risk Management: The Ethical Considerations. *Jefferson Journal of Psychiatry*, *11*(2). https://doi.org/10.29046/JJP.011.2.006

Clouser, K. D., Culver, C. M., & Gert, B. (1997). Malady. In J. M. Humber & R. F. Almeder (Eds.), *What is disease?* (pp. 173–217). Humana Press. https://doi.org/10.1007/978-1-59259-451-1_3

Cooper, D. (1967). *Psychiatry and anti-psychiatry*. Tavistock Publications.

Cooper, R. V. (2002). Disease. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *33*(2), 263–282. https://doi.org/10.1016/S0039-3681(02)00018-3

Cooper, R. V. (2020). The concept of disorder revisited: Robustly value-laden despite change. *Aristotelian Society Supplementary Volume*, *94*(1), 141–161. https://doi.org/10.1093/arisup/akaa010

Corrigan, P. W. (2000). Mental health stigma as social attribution: Implications for research methods and attitude change. *Clinical Psychology: Science and Practice*, *7*(1), 48–67. https://doi.org/10.1093/clipsy.7.1.48

Corrigan, P. W., & Miller, F. (2004). Shame, blame, and contamination: A review of the impact of mental illness stigma on family members. *Journal of Mental Health*, *13*(6). https://www.tandfonline.com/doi/abs/10.1080/09638230400017004?casa_token=F9nkpGFp3NcAAAAA:FWi960Uc5cI7M_IJeLI2TEJsuITa9LhR1Szz_wFzF2ipNgy0zG7AAEkyRMhz0gFQogOMrnud4dJ8Cnc

Craigie, J. (2015). Against a singular understanding of legal capacity: Criminal responsibility and the Convention on the Rights of Persons with Disabilities. *International Journal of Law and Psychiatry*, *40*, 6–14. https://doi.org/10.1016/j.ijlp.2015.04.002

*Daughter*. (2022, May 20). Https://Dictionary.Cambridge.Org/Dictionary/English/Daughter. https://dictionary.cambridge.org/dictionary/english/daughter

Dean, J., & Keshavan, M. (2017). The neurobiology of depression: An integrated view. *Asian Journal of Psychiatry*, *27*, 101–111. https://doi.org/10.1016/j.ajp.2017.01.025

*Debate Over Gaming Disorder Is Not All Fun and Games | Psychology Today*. (n.d.). Retrieved October 11, 2021, from https://www.psychologytoday.com/intl/blog/making-meaning/201901/debate-over-gaming-disorder-is-not-all-fun-and-games

Deusenbery, C., Wang, Y., & Shukla, A. (2021). Recent Innovations in Bacterial Infection Detection and Treatment. *ACS Infectious Diseases*, *7*(4), 695–720. https://doi.org/10.1021/acsinfecdis.0c00890

Drescher, J. (2015). Out of DSM: Depathologizing Homosexuality. *Behavioral Sciences*, *5*(4), 565–575. https://doi.org/10.3390/bs5040565

Dutilh Novaes, C. (2018). Carnapian explication and ameliorative analysis: A systematic comparison. *Synthese*. https://doi.org/10.1007/s11229-018-1732-9

Dutilh Novaes, C. (2020). Carnapian explication and ameliorative analysis: A systematic comparison. *Synthese*, *197*(3), 1011–1034. https://doi.org/10.1007/s11229-018-1732-9

Dworkin, G. (2020). Paternalism. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/fall2020/entries/paternalism/

Ehrenreich, B., & English, D. (2013). *For her own good: Two centuries of the experts advice to women*. Knopf Doubleday Publishing Group.

Engelhardt, H. T. (1974). The Disease of Masturbation: Values and the Concept of Disease. *Bulletin of the History of Medicine*, *48*(2), 234–248.

Engelhardt, H. T., Jr. (1976). Ideology and Etiology. *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, *1*(3), 256–268. https://doi.org/10.1093/jmp/1.3.256

Ereshefsky, M. (2009). Defining 'health' and 'disease.' *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *40*(3), 221–227. https://doi.org/10.1016/j.shpsc.2009.06.005

Erol, A. (2017). How to Conduct Scientific Research? *Archives of Neuropsychiatry*, *54*(2), 97–98. https://doi.org/10.5152/npa.2017.0120102

Feit, N. (2017). Harm and the concept of medical disorder. *Theoretical Medicine and Bioethics*, *38*(5), 367–385. https://doi.org/10.1007/s11017-017-9418-8

Fletcher, G. (2016). *The Philosophy of Well-Being: An Introduction*. Routledge. https://doi.org/10.4324/9781315745329

Forest, D., & Le Bidan, M. (2016). In search of normal functions: BST, Cummins functions, and Hempel's problem. In É. Giroux (Ed.), *Naturalism in the Philosophy of Health* (pp. 39–51). Springer International Publishing. https://doi.org/10.1007/978-3-319-29091-1_3

Foucault, M. (1971). *Madness and Civilization*. Routledge. https://doi.org/10.4324/9780203164693

Fredrickson, B. L., & Joiner, T. (2002). Positive Emotions Trigger Upward Spirals Toward Emotional Well-Being. *Psychological Science*, *13*(2), 172–175. https://doi.org/10.1111/1467-9280.00431

Fulford, K. W. M. (1989). *Moral theory and medical practice*. Cambridge University Press.

Garson, J. (2012). Function, selection, and construction in the brain. *Synthese*, *189*(3), 451–481.

Garson, J. (2016). *A critical overview of biological functions*. Springer.

Gettier, E. (1963). Is knowledge justified true belief? *Analysis*, *23*(6), 121–123.

Glover, J. (1970). *Responsibility*. Humanities P.

Goffman, E. (1961). *Asylums: Essays on the social situation of mental patients and other inmates*. Anchor Books.

Goffman, E. (2009). *Stigma: Notes on the management of spoiled identity*. Simon and Schuster.

Graham, G. (2013). *The disordered mind: An introduction to philosophy of mind and mental illness* (Second edition). Routledge.

Grđan, K. (2010). *Legal Status of Sexual and Gender Minorities in South-eastern Europe* (p. 21). https://doi.org/10.13140/RG.2.1.4903.0489

Grob, G. N. (2011). The attack of psychiatric legitimacy in the 1960s: Rhetoric and reality. *Journal of the History of the Behavioral Sciences*, *47*(4), 398–416. https://doi.org/10.1002/jhbs.20518

Hacking, I. (1996). The looping effects of human kinds. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition* (pp. 351–383). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198524021.003.0012

Hacking, I. (1999). *The social construction of what?* Harvard University Press.

Hall, T. (1975). *History of General Physiology: 600B.C to A.D. 1900, Vol. I from Pre-Socratic Times to the Enlightenment*. University of Chicago Press. https://philpapers.org/rec/HALHOG-2

Hamber, B., & Brian, R. (1996). *Mental Illness and Human Rights: In Search of Context, Consequence and Effective Care*. *1*(1), 72–91.

Harrington, A. (2019). *Mind Fixers: Psychiatry's Troubled Search for the Biology of Mental Illness*. W. W. Norton & Company.

Hart, H. L. A. (2008). *Punishment and responsibility: Essays in the philosophy of law* (2nd Revised edition). Oxford University Press.

Hawthorne, J. (2021). Inductive Logic. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/logic-inductive/

Herek, G. M., & Garnets, L. (2015). Hooker, Evelyn (1907–1996). In *The International Encyclopedia of Human Sexuality* (pp. 501–581). American Cancer Society. https://doi.org/10.1002/9781118896877.wbiehs219

Hesslow, G. (1993). Do we need a concept of disease? *Theoretical Medicine*, *14*(1), 1–14. https://doi.org/10.1007/BF00993984

Hobfoll, S. E. (2002). Social and Psychological Resources and Adaptation. *Review of General Psychology*, *6*(4), 307–324. https://doi.org/10.1037/1089-2680.6.4.307

Hofmann, B. (2001). Complexity of the Concept of Disease as Shown Through Rival Theoretical Frameworks. *Theoretical Medicine and Bioethics*, *22*(3), 211–236. https://doi.org/10.1023/A:1011416302494

Hong, L. K. (1984). Survival of the fastest: On the origin of premature ejaculation. *The Journal of Sex Research*, *20*(2), 109–122. https://doi.org/10.1080/00224498409551212

Hooker, E. (1956). A preliminary analysis of group behavior of homosexuals. *The Journal of Psychology*, *42*(2), 217–225. https://doi.org/10.1080/00223980.1956.9713035

Hooker, E. (1957). The adjustment of the male overt homosexual. *Journal of Projective Techniques*, *21*(1), 18–31. https://doi.org/10.1080/08853126.1957.10380742

Hooker, E. (1993). Reflections of a 40-year exploration: A scientific view on homosexuality. *American Psychologist*, *48*(4), 450–453. https://doi.org/10.1037/0003-066X.48.4.450

Horwitz, A. V., & Scheid, T. L. (1999). *A Handbook for the Study of Mental Health: Social Contexts, Theories, and Systems*. Cambridge University Press.

Hubbard, E. M. (2007). Neurophysiology of synesthesia. *Current Psychiatry Reports*, *9*(3), 193–199. https://doi.org/10.1007/s11920-007-0018-6

Jhee, J. H., Lye, Q. T., & Woo, K. (2019). Not Yet Game Over: A Reappraisal of Video Game Addiction. *Loading: The Journal of the Canadian Game Studies Association*, *12*(19). https://doi.org/10.7202/1058317ar

Jurjako, M. (2011). Luca Malatesti and John McMillan (Eds.), Responsibility and Psychopathy: Interfacing Law, Psychiatry, and Philosophy. *Prolegomena*, *10*(1), 147–154.

Jurjako, M. (2019). Is psychopathy a harmful dysfunction? *Biology & Philosophy*, *34*(5). https://doi.org/10.1007/s10539-018-9668-5

Jurjako, M., & Malatesti, L. (2018). Neuropsychology and the criminal responsibility of psychopaths: Reconsidering the evidence. *Erkenntnis*, *83*(5), 1003–1025. https://doi.org/10.1007/s10670-017-9924-0

Kaplan, M. (1983). A woman's view of DSM-III. *American Psychologist*, *38*(7), 786–792. https://doi.org/10.1037/0003-066X.38.7.786

Kass, L. R. (1975). Regarding the end of medicine and the pursuit of health. *The Public Interest*, *40*, Unknown.

Kendell, R. (1975). The concept of disease and its implications for psychiatry. *The British Journal of Psychiatry : The Journal of Mental Science*. https://doi.org/10.1192/BJP.127.4.305

Kincaid, H. (2008). Do we need theory to study disease? Lessons from cancer research and their implications for mental illness. *Perspectives in Biology and Medicine*, *51*(3), 367–378. https://doi.org/10.1353/pbm.0.0019

Kingma, E. (2007). What is it to be healthy? *Analysis*, *67*(294), 128–133.

Kingma, E. (2014). Naturalism about health and disease: Adding nuance for progress. *The Journal of Medicine and Philosophy*, *39*(6), 590–608. https://doi.org/10.1093/jmp/jhu037

Lachner, C., Maniaci, M. J., Vadeboncoeur, T. F., Dawson, N. L., Rummans, T. A., Roy, A., Hall, L. L., & Burton, M. C. (2020). Are pre-existing psychiatric disorders the only reason for involuntary holds in the emergency department? *International Journal of Emergency Medicine*, *13*(1), 4. https://doi.org/10.1186/s12245-020-0265-4

LaPorte, J. (2009). *Natural Kinds and Conceptual Change*. Cambridge University Press.

Lemoine, M. (2013). Defining disease beyond conceptual analysis: An analysis of conceptual analysis in philosophy of medicine. *Theoretical Medicine and Bioethics*, *34*(4), 309–325. https://doi.org/10.1007/s11017-013-9261-5

Ludlow, P. (2014). *Living words: Meaning underdetermination and the dynamic lexicon*. OUP Oxford.

*Maladaptive Behaviors definition*. (n.d.). [Psychology Glossary]. AlleyDog.Com. Retrieved September 16, 2022, from https://www.alleydog.com/glossary/definition.php?term=Maladaptive+Behaviors

Malatesti, L., Jurjako, M., & Meynen, G. (2020). The insanity defence without mental illness? Some considerations. *International Journal of Law and Psychiatry*, *71*, 101571. https://doi.org/10.1016/j.ijlp.2020.101571

Margolis, E., & Laurence, S. (2021). Concepts. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2021). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/spr2021/entries/concepts/

McLean, I., & McMillan, A. (2009). The Concise Oxford Dictionary of Politics. In *The Concise Oxford Dictionary of Politics*. Oxford University Press. https://www.oxfordreference.com/view/10.1093/acref/9780199207800.001.0001/acref-9780199207800

McRae, A., & Human Rights Watch (Eds.). (2010). *"Once you enter, you never leave": Deinstitutionalization of persons with intellectual or mental disabilities in Croatia*. Human Rights Watch.

Miller, F. G., Brody, H., & Chung, K. C. (2000). Cosmetic surgery and the internal morality of medicine. *Cambridge Quarterly of Healthcare Ethics: CQ: The International Journal of Healthcare Ethics Committees*, *9*(3), 353–364. https://doi.org/10.1017/s0963180100903074

Mohan, H. (2015). *Textbook of Pathology* (7th ed.). Jaypee Brothers Medical Publishers (P) Ltd. http://172.16.8.253:7003/jspui/handle/123456789/540

Moseley, D. D., & Gala, G. (2015). *Philosophy and Psychiatry: Problems, Intersections and New Perspectives*. Routledge.

Moskop, J. C. (1999). Informed Consent in the Emergency Department. *Emergency Medicine Clinics of North America*, *17*(2), 327–340. https://doi.org/10.1016/S0733-8627(05)70062-6

Muckler, D., & Taylor, J. S. (2020). The irrelevance of harm for a theory of disease. *The Journal of Medicine and Philosophy*, *45*(3), 332–349. https://doi.org/10.1093/jmp/jhaa007

Murphy, D. (2005). The Concept of Mental Illness—Where the Debate has Reached and Where it Needs to Go. *Journal of Theoretical and Philosophical Psychology*, *25*(1), 116–132. https://doi.org/10.1037/h0091255

Murphy, D. (2006). *Psychiatry in the scientific image*. The MIT Press.

Murphy, D., & Woolfolk, R. L. (2000). The harmful dysfunction analysis of mental disorder. *Philosophy, Psychiatry, & Psychology*, *7*(4), 241–252.

Nesse, R. M., & Stein, D. J. (2012). Towards a genuinely medical model for psychiatric nosology. *BMC Medicine*, *10*(1), 5. https://doi.org/10.1186/1741-7015-10-5

Nordby, H. (2016). The analytic–synthetic distinction and conceptual analyses of basic health concepts. *Medicine, Health Care and Philosophy*, *9*, 169–180. https://doi.org/10.1007/s11019-006-0002-7

Nordenfelt, L. (1993). On the relevance and importance of the notion of disease. *Theoretical Medicine*, *14*(1), 15–26. https://doi.org/10.1007/BF00993985

Nordenfelt, L. (1995). *On the nature of health: An action-theoretic approach*. Springer Science & Business Media.

Nordenfelt, L. (2007). The concepts of health and illness revisited. *Medicine, Health Care, and Philosophy*, *10*(1), 5–10. https://doi.org/10.1007/s11019-006-9017-3

Overgaard, S., Gilbert, P., & Burwood, S. (2013). *An Introduction to Metaphilosophy*. Cambridge University Press.

Parfit, D. (1984). *Reasons and Persons*. OUP Oxford.

Parikh, R., & Renero, A. (2017). Justified True Belief: Plato, Gettier, and Turing. In J. Floyd & A. Bokulich (Eds.), *Philosophical Explorations of the Legacy of Alan*

*Turing: Turing 100* (pp. 93–102). Springer International Publishing. https://doi.org/10.1007/978-3-319-53280-6_4

Peterson, J. W. (1996). Bacterial Pathogenesis. In S. Baron (Ed.), *Medical Microbiology* (4th ed.). University of Texas Medical Branch at Galveston. http://www.ncbi.nlm.nih.gov/books/NBK8526/

Phelan, J. C. (2005). Geneticization of deviant behavior and consequences for stigma: The case of mental illness. *Journal of Health and Social Behavior*, *46*(4), 307–322. https://doi.org/10.1177/002214650504600401

Phillips, M., R. (2009). Is distress a symptom of mental disorders, a marker of impairment, both or neither? *World Psychiatry*, *8*(2), 91–92.

Plato. (1997). *Plato: Complete Works*. Hackett Publishing.

Quine, W. van O. (1976). Two Dogmas of Empiricism. In S. G. Harding (Ed.), *Can Theories be Refuted? Essays on the Duhem-Quine Thesis* (pp. 41–64). Springer Netherlands. https://doi.org/10.1007/978-94-010-1863-0_2

Quine, W. van O. (2013). *Word and object*. MIT Press.

Räikkä, J. (1996). The social concept of disease. *Theoretical Medicine*, *17*(4), 353–361. https://doi.org/10.1007/BF00489680

Reznek, L. (1987). *The Nature of Disease*. Routledge & Kegan Paul.

Ritsher, J. B., Otilingam, P. G., & Grajales, M. (2003). Internalized stigma of mental illness: Psychometric properties of a new measure. *Psychiatry Research*, *121*(1), 31–49. https://doi.org/10.1016/j.psychres.2003.08.008

Rowland, D. L., & Cooper, S. E. (2022). The Tenuous Role of Distress in the Diagnosis of Premature Ejaculation: A Narrative Review. *Sexual Medicine*, *10*(5), 100546. https://doi.org/10.1016/j.esxm.2022.100546

Russell, B. (2020). A Priori Justification and Knowledge. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2020/entries/apriori/

Russell, L. B. (1992). Opportunity costs in modern medicine. *Health Affairs*, *11*(2), 162–169. https://doi.org/10.1377/hlthaff.11.2.162

Scadding, J. G. (1988). Health and disease: What can medicine do for philosophy? *Journal of Medical Ethics*, *14*(3), 118–124. https://doi.org/10.1136/jme.14.3.118

Schwartz, P. H. (2007a). Decision and discovery in defining disease. In H. Kincaid & J. McKitrick (Eds.), *Establishing Medical Reality: Essays in the Metaphysics and the Epistemology of Biomedical Science* (pp. 47–63). Springer.

Schwartz, P. H. (2007b). Defining dysfunction: Natural selection, design, and drawing a line. *Philosophy of Science*, *74*(3), 364–385. https://doi.org/10.1086/521970

Schwartz, P. H. (2014). Reframing the disease debate and defending the biostatistical theory. *The Journal of Medicine and Philosophy*, *39*(6), 572–589. https://doi.org/10.1093/jmp/jhu039

Shiffrin, S. valentine. (2012). Harm and its moral significance. *Legal Theory*, *18*(3), [i]-398.

Sompayrac, L. (2004). *How Cancer Works*. Jones & Bartlett Learning.

Sparrow, R. (2005). Defending Deaf Culture: The Case of Cochlear Implants*. *Journal of Political Philosophy*, *13*(2), 135–152. https://doi.org/10.1111/j.1467-9760.2005.00217.x

Spitzer, R. L. (1981). The diagnostic status of homosexuality in DSM-III: A reformulation of the issues. *The American Journal of Psychiatry*, *138*(2), 210–215. https://doi.org/10.1176/ajp.138.2.210

Spitzer, R. L., & Wakefield, J. C. (1999). DSM-IV diagnostic criterion for clinical significance: Does it help solve the false positives problem? *American Journal of Psychiatry*, *156*(12), 1856–1864. https://doi.org/10.1176/ajp.156.12.1856

Stegenga, J. (2018). *Care and cure: An introduction to philosophy of medicine*. University of Chicago Press.

Strawson, P. F. (1992). Analysis and Metaphysics: An Introduction to Philosophy. In *Analysis and Metaphysics*. Oxford University Press. https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780198751182.001.0001/acprof-9780198751182

Szasz, T. S. (1960). The myth of mental illness. *American Psychologist*, *15*(2), 113–118. https://doi.org/10.1037/h0046535

Tekin, S. (2012, June 12). *Theorizing Looping Effects: Lessons from Cognitive Sciences*. http://philsci-archive.pitt.edu/9332/

Terry, P. B. (2007). Informed Consent in Clinical Medicine. *Chest*, *131*(2), 563–568. https://doi.org/10.1378/chest.06-1955

*The History of Hysteria*. (n.d.). Office for Science and Society. Retrieved November 2, 2022, from https://www.mcgill.ca/oss/article/history-quackery/history-hysteria

Thornton, T. (2007). *Essential Philosophy of Psychiatry*. OUP Oxford.

Tiberius, V., & Plakias, A. (2010). Well-being. In *The moral psychology handbook* (pp. 402–432). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199582143.003.0013

Trautmann, S., Rehm, J., & Wittchen, H. (2016). The economic costs of mental disorders. *EMBO Reports*, *17*(9), 1245–1249. https://doi.org/10.15252/embr.201642951

Turner, A. J. (2010). Are disorders sufficient for reduced responsibility? *Neuroethics*, *3*(2), 151–160. https://doi.org/10.1007/s12152-009-9041-0

Wahl, O. F. (2012). Stigma as a barrier to recovery from mental illness. *Trends in Cognitive Sciences*, *16*(1), 9–10. https://doi.org/10.1016/j.tics.2011.11.002

Wakefield, J. C. (1992). The concept of mental disorder. On the boundary between biological facts and social values. *The American Psychologist*, *47*(3), 373–388.

Wakefield, J. C. (1997a). Normal inability versus pathological disability: Why Ossorio's definition of mental disorder is not sufficient. *Clinical Psychology: Science and Practice*, *4*(3), 249–258. https://doi.org/10.1111/j.1468-2850.1997.tb00113.x

Wakefield, J. C. (1997b). When is development disordered? Developmental psychopathology and the harmful dysfunction analysis of mental disorder. *Development and Psychopathology*, *9*(2), 269–290. https://doi.org/10.1017/S0954579497002058

Wakefield, J. C. (1997c). Diagnosing DSM-IV—Part I: DSM-IV and the concept of disorder. *Behaviour Research and Therapy*, *35*(7), 633–649. https://doi.org/10.1016/S0005-7967(97)00018-1

Wakefield, J. C. (2006). What makes a mental disorder mental? *Philosophy, Psychiatry, & Psychology*, *13*(2), 123–131. https://doi.org/10.1353/ppp.2007.0010

Wakefield, J. C. (2007). The concept of mental disorder: Diagnostic implications of the harmful dysfunction analysis. *World Psychiatry*, *6*(3), 149–156.

Wakefield, J. C. (2011). DSM-5 proposed diagnostic criteria for sexual paraphilias: Tensions between diagnostic validity and forensic utility. *International Journal of Law and Psychiatry*, *34*(3), 195–209. https://doi.org/10.1016/j.ijlp.2011.04.012

Wakefield, J. C. (2014). The biostatistical theory versus the harmful dysfunction analysis, part 1: Is part-dysfunction a sufficient condition for medical disorder? *Journal of Medicine and Philosophy*, *39*(6), 648–682. https://doi.org/10.1093/jmp/jhu038

Wakefield, J. C., & Conrad, J. A. (2019). Does the harm component of the harmful dysfunction analysis need rethinking?: Reply to Powell and Scarffe. *Journal of Medical Ethics*, *45*(9), 594–596. https://doi.org/10.1136/medethics-2019-105578

Wakefield, J. C., & Conrad, J. A. (2020). Harm as a necessary component of the concept of medical disorder: Reply to Muckler and Taylor. *The Journal of Medicine and Philosophy*, *45*(3), 350–370. https://doi.org/10.1093/jmp/jhaa008

Walker, M. J., & Rogers, W. A. (2018). A new approach to defining disease. *The Journal of Medicine and Philosophy*, *43*(4), 402–420. https://doi.org/10.1093/jmp/jhy014

Wasserman, D., & Aas, S. (2022). Disability: Definitions and Models. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2022). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/archives/sum2022/entries/disability/

*WHO expert defends gaming disorder listing: "This moral panic lives its own life."* (n.d.). GamesIndustry.Biz. Retrieved October 11, 2021, from

https://www.gamesindustry.biz/articles/2018-06-26-who-expert-defend-gaming-disorder-listing-this-moral-panic-lives-its-own-life

World Health Organization. (1980). *International classification of impairments, disabilities, and handicaps: A manual of classification relating to the consequences of disease, published in accordance with resolution WHA29.35 of the Twenty-ninth World Health Assembly, May 1976*. World Health Organization. https://apps.who.int/iris/handle/10665/41003

World Health Organization (Ed.). (1992). *The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines*. World Health Organization.

Worrall, J., & Worrall, J. (2001). Defining Disease: Much Ado About Nothing? In A.-T. Tymieniecka & E. Agazzi (Eds.), *Life Interpretation and the Sense of Illness Within the Human Condition* (pp. 33–55). Kluwer Academic Publishers.