

Georgia State University

ScholarWorks @ Georgia State University

Philosophy Theses

Department of Philosophy

8-10-2021

What Can an Evolutionary Explanation Bring to The Demarcation of The Normal from The Pathological in Psychiatry? Nesse's Case of Depression.

Chrysi Malouchou Kanellopoulou

Follow this and additional works at: https://scholarworks.gsu.edu/philosophy_theses

Recommended Citation

Malouchou Kanellopoulou, Chrysi, "What Can an Evolutionary Explanation Bring to The Demarcation of The Normal from The Pathological in Psychiatry? Nesse's Case of Depression.." Thesis, Georgia State University, 2021.

https://scholarworks.gsu.edu/philosophy_theses/303

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact scholarworks@gsu.edu.

What Can an Evolutionary Explanation Bring to The Demarcation of The Normal from The
Pathological in Psychiatry? Nesse's Case of Depression.

by

Chrysi Maria Malouchou Kanellopoulou

Under the Direction of Andrea Scarantino, PhD and Sebastian Rand, PhD

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2021

ABSTRACT

Randolph Nesse argues that evolutionary theory is the key element in elaborating a valid criterion demarcating the normal from the pathological in psychiatry (Nesse, 2001, 2009, 2015, 2017). By focusing on the application of Nesse's criterion on the demarcation of normal low mood from pathological depression, I argue – contrary to Nesse's claims – that evolutionary theory cannot generate a valid criterion from the differentiation of normal low mood states from pathological depression. Indeed, expression in conformity to evolved functions cannot constitute a sufficient condition for normality, as Nesse's project should imply. Moreover, grounding normality in evolved functions presupposes a kind of fixity of our emotional states, which seems unwarranted in the light of a constantly changing environment. As a result of these limitations, I suggest that the relevant distinction in psychiatry should not be “normality” *versus* “abnormality”, but, rather, “that which requires intervention” *versus* “that which does not”.

INDEX WORDS: Depression, Evolution, Normal, Pathological, Functions, Nesse, Psychiatry

Copyright by
Chrysi Maria Malouchou Kanellopoulou
2021

What Can an Evolutionary Explanation Bring to The Demarcation of The Normal from The
Pathological in Psychiatry? Nesse's Case of Depression

by

Chrysi Maria Malouchou Kanellopoulou

Committee Chairs: Andrea Scarantino
Sebastian Rand

Electronic Version Approved:

Office of Graduate Services
College of Arts and Sciences
Georgia State University
May 2021

ACKNOWLEDGEMENTS

First, I would like to express my profound gratitude to my advisors, Dr. Andrea Scarantino and Dr. Sebastian Rand. Dr. Scarantino invited me to challenge and clarify my reasoning: he encouraged me to keep pushing my thoughts further, even when I believed I had exhausted them. Dr. Rand helped me engage with the thesis under a new perspective, which revealed stimulating new ideas. There are no words to describe how much I have learned with them through writing this thesis and through their classes, both as a student and as a person. Their encouragement and support throughout this program have been invaluable. I cannot thank them enough.

I also owe a lot to the entire philosophy department. To the faculty, who has given me the unique opportunity to be a part of this program, and my fellow classmates for their wonderful and insightful discussions. I would especially like to thank Alex for his feedback on my draft and Theodore for the illuminating discussions on depression.

I also want to thank Elisabetta for her perceptive comments on evolutionary theory, and Sherin for sharing her thorough knowledge on depression. Finally, I would like to wholeheartedly thank my friends, Nicolas, and my family for always being there for me. I could not have done this thesis without them.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS		IV
1	INTRODUCTION	1
2	THE DEMARCATION PROBLEM IN PSYCHIATRY AND NESSE'S PROPOSAL	3
2.1	A Depression Pseudo-epidemic?	3
2.2	Nesse's Proposal	8
3	CAN EVOLUTIONARY THEORY GENERATE A VALID MODEL TO DEMARCATATE THE NORMAL FROM THE PATHOLOGICAL?	15
3.1	Is Low Mood's Conformity with its Etiological Function a Sufficient Condition for its Normality?	18
3.2	Can we Ground Normality in Etiological Functions?	23
<i>3.2.1. Are Non-selected States Pathologies?</i>		<i>24</i>
<i>3.2.1. A Discrepancy between Evolutionary Functions and Causal Role Functions?....</i>		<i>26</i>
4	A POTENTIAL SOLUTION: REFORMULATING THE NORMAL-ABNORMAL DISTINCTION IN PSYCHIATRY?	30
4.1	Abandoning the Functionalist Framework?	30
4.2	Normality-Abnormality: an Irrelevant Distinction?	33
4.3.	An Intervention-based Approach	35
5	CONCLUSION	39
	REFERENCES	40

1 INTRODUCTION

How can a clinician know whether their patient is suffering from a depressive pathology or whether they are merely experiencing a set of uncomfortable but normal emotional reactions? Should judgments about the pathological nature of their emotional state be derived solely from considerations on the intensity and duration of their symptoms or from an evaluation of the context in which the symptoms have occurred? Unfortunately, psychiatry lacks a straightforward and uncontroversial answer to these questions: there is no agreement on how the demarcation of the normal from the pathological should be articulated, and hence, no consensus on the criteria a clinician should follow in their diagnoses.

In a series of papers, book chapters and monographs, Randolph Nesse (2001, 2009, 2015, 2017) argues that evolutionary theory can provide psychiatry the required tools to elaborate an objective and valid demarcation criterion for emotional disorders: the normality of an emotional state should be assessed based on the state's conformity with its evolved function; if the state is expressed in conformity with its function, it is normal. If not, it is pathological.

This thesis aims at evaluating Nesse's project of using evolutionary theory to generate a criterion for the demarcation of the normal from the pathological in psychiatry, specifically for depression: can evolutionary theory provide the solution to the demarcation problem for depression? After laying out Nesse's position in the first section, the second section will argue that there are essential limitations to the project of appealing to evolutionary functions to determine the normality or abnormality of a depressed state. First, Nesse's own understanding of pathological depression reveals not only that there is more to the pathological than a mere lack of conformity with functional norms, but also that satisfaction of evolved functions cannot constitute a sufficient condition for normality, hence undermining Nesse's project's scope.

Secondly, I will show that the justification for grounding normality in evolved functions is dubious. It presupposes that our emotional states are currently triggered and expressed in complete conformity to the way they were in the past, at the time they were selected for. This kind of fixity of our emotional states seems unjustified in the light of a constantly changing environment. Finally, in the third section, I will suggest a different approach to the demarcation problem, in the attempt of overcoming these limitations: the idea is to replace the distinction “normal-abnormal” with “that which requires intervention-that which does not” as the relevant one for psychiatric diagnosis.

2 THE DEMARCATION PROBLEM IN PSYCHIATRY AND NESSE'S PROPOSAL

2.1 A Depression Pseudo-epidemic?

According to the World Health Organization, more than 264 million people worldwide are suffering from depression,¹ whether it is a mild, moderate, or severe form of depression, chronic or not. High rates of depression are shown to have significantly increased in the past decades to the point that many researchers are speaking of a “depression epidemic” (Hidaka, 2012). That tendency is predicted to go on: projections indicate that by 2030 depression will be the “single leading cause of disease... in high income countries” (Ormel et al. 2020, 1).

The explanation behind the worldwide increased prevalence of depression is not, however, straightforward. As Wakefield and Demazeux note, two kinds of rival explanations can be pursued (Wakefield and Demazeux, 2016, 6). According to the first kind, there is a real increase in depression cases in the world. Humans are increasingly becoming pathologically depressed, due to social and environmental changes, such as changes in dietary habits, “the rise of individualism, the development of neoliberalism, the culture of narcissism, social mobility, [or] constant exposure to images of those with greater beauty or wealth through the media” (idem, 6). According to the second kind of explanation, the increase in the prevalence of depression does not correspond to a real increase in pathological cases. The increase is explained both by the unreliability and lack of sufficient studies that aimed at measuring the prevalence of depression until the second half of the 20th century – which resulted in an *underestimation* of pathological cases – but also by the lack of validity of the diagnostic criteria currently used, which are resulting in an *overestimation* of pathological cases. According to this kind of explanation, the depression epidemic is, in fact, a “pseudo-epidemic” (Summerfield, 2006, 161):

¹ WHO, <https://www.who.int/news-room/fact-sheets/detail/depression>, 2021.

the criteria commonly used to diagnose depression are too broad and ultimately conflate so-called “normal” emotional reactions, emerging as ordinary responses to adverse circumstances, with pathological depression (Wakefield and Demazeux, 2016, 6).

The *Diagnostic and Statistical Manual of Mental Disorders* (DSM), currently in its fifth version, (DSM-5) published by the American Psychiatric Association (APA, 2013), is the main target of the pseudo-epidemic explanation: widely used in clinical settings, statistical demographic studies (First et al., 2014; Hasin et al., 2018), as well as research programs probing potential correlations between depression and biological or neurological mechanisms (Doyle et al., 2019; Burkhouse et al., 2017), it is seen as the primary driver of the pathologization of normal emotional reactions, such as sadness, grief, or low mood (Wakefield and Demazeux, 2016).

In fact, starting with its third version (APA, 1980), the DSM has been committed to an atheoretical approach to the classification and diagnosis of mental disorders. The lack of commonly accepted explanations of the nature of most mental disorders or of their precise etiology, as well as the absence of disorder-specific net biological markers for diagnostic tests,² has led psychiatrists to favor a diagnostic and classification system that relies solely on the presence and intensity of symptoms: conflicting theories could result in conflicting diagnoses and classifications, thus rendering communication and research across psychiatrists inconsistent (Nesse and Stein, 2012). As Nesse describes it, the DSM pursues a “core strategy of using checklists of criteria to define diagnostic categories” (idem, 2). In the case of depression, the DSM-5 states that an individual can be diagnosed with major depressive disorder if they exhibit, for at least two consecutive weeks, “either depressed mood or the loss of interest or pleasure in

² Biological markers have not been identified for most disorders in the DSM, “with the exception of neurological disorders such as Huntington’s disease” (Nesse and Stein, 2012, 2).

nearly all activities” (APA, 2013, 163), as well as four other symptoms from the following list: significant loss or gain of weight, decrease or increase in appetite, a slowing down of thought or physical activity, fatigue or loss of energy, feelings of worthlessness or guilt, diminished ability to concentrate and recurrent thoughts of death or suicidal attempts (APA, 2013, 160-161).³

This atheoretical approach has two significant consequences for the diagnosis of depression. First, the assessment of the symptoms is context-independent: no matter the circumstances in which the symptoms have appeared, a person can be diagnosed with major depressive disorder as long as their symptoms satisfy the checklist. Second, given that the distinction between an abnormal and a normal state is not derived from any theoretical framework, the validity of the two-week time threshold is not guaranteed (Nesse and Jackson, 2011): there is no genuine explanation supporting that after two weeks – a relatively low threshold – the symptoms indicate a pathology.

Nesse argues that these consequences of the DSM’s atheoretical approach are the main reason why the DSM is seen as encouraging pathologization of normal emotional reactions. If clinicians judge whether a pathology is present by relying on a low time-threshold – which does not guarantee the pathology’s presence – and ignore the context in which the symptoms arise, they are inclined to put less weight on the possibility that their patients are undergoing a normal emotional reaction to complex circumstances (idem, 2011). Suppose, for instance, that X experiences symptoms of bereavement after the loss of their life partner, such as changes in

³ Six different disorders are classified under the “Depressive Disorders” category in the DSM-5: “disruptive mood dysregulation disorder, major depressive disorder..., persistent depressive disorder (dysthymia), premenstrual dysphoric disorder, substance/medication-induced disorder, depressive disorder due to another medical condition”, as well as “other specified [or] unspecified depressive disorder[s]” (APA, 2013, 155). What is usually meant when talking about depression is major depressive disorder. Disruptive mood dysregulation disorder is specific to children; persistent depressive disorder denotes a chronic – at least two year long – major depressive disorder. Premenstrual dysphoric disorder, substance abuse disorder or depressive disorders due to another medical condition, all map, to a certain extent, onto major depressive disorder but are specific to certain chemical changes in one’s organism.

eating patterns, loss of motivation, fatigue and general low mood for a period of merely two weeks. According to the DSM-5, X can qualify as being pathologically depressed. This conclusion is at odds with both our commonsense intuitions and with the view supported by many experts that these symptoms arise as a normal reaction to loss; that they do not necessarily indicate the presence of a pathology (Maj, 2016).

It is important to note that not all practitioners follow the DSM in their diagnosis, and if they do, they do not necessarily follow it to the letter (First et al., 2014). Many clinicians might be hesitant to diagnose a patient who has just lost their life partner with major depressive disorder. In a note, the DSM-5 also invites to the “exercise of critical judgment based on the individual’s history and the cultural norms for the expression of distress in the context of loss” (APA, 2013, 161). However, despite this note, it seems clear that the DSM’s stance towards such cases is to ignore the particularity of the context and encourage the diagnosis of a disorder. In fact, while earlier versions of the DSM included an explicit bereavement exclusion clause – mentioning that a patient should not be diagnosed with major depressive disorder if they experience depression-like symptoms up to two months after the loss of a loved one – this clause was removed in the DSM-5 (Castiglione and Laudisa, 2015). One important argument in favor of its removal is that there is no reliable way for the practitioner to know that the individual is *not* suffering from a pathology, not only because symptoms of bereavement and of major depression overlap, but also because the circumstances might have precipitated or even caused a state of pathological depression (Kendler, Myers and Zisook, 2008). Given that not receiving treatment if a patient needs one is riskier than providing a treatment when not needed, if in doubt, the practitioner’s safer choice could be to diagnose them with major depressive disorder (Zisook and

Shuchter, 1991). Hence, while there is an important justification behind the clause's removal, it is the reason why many view it as favoring the misidentification of normal grief with pathology.⁴

Nesse aims precisely at avoiding the DSM's issues by presenting a theoretical alternative to the diagnosis and classification of depression – along with all other psychiatric disorders. Based on evolutionary theory and functional explanations of moods and emotions, he develops a research project that seeks to provide a “properly scientific” foundation for psychiatry from which “objective” criteria demarcating the normal from the pathological could be generated (Nesse and Stein, 2012; Nesse, 2017). Finding a reliable and valid way to demarcate the normal from the pathological is crucial for individuals to receive the right treatment for their situation, for research programs inquiring its causes or the biological and neurological mechanisms underlying it, as well as for public health policies (see Horwitz and Wakefield, 2007, for the potential risks of unwarranted demarcation criteria).

The question that I will try to answer in this thesis is hence the following: can Nesse's evolutionary explanation of depression generate a valid model for the demarcation of normal emotional reactions from pathological depression? I will try to show that, while integrating evolutionary biology to psychiatry can provide a deeper understanding of psychiatric disorders by shedding light on their origins, it cannot generate a genuinely scientific demarcation criterion, as Nesse claims. More specifically, the validity of a model demarcating normal emotional states from pathological depression by relying on the former's evolutionary function is challenged by the observation of an internal inconsistency in Nesse's model, revealing that a state's expression in accordance with its evolutionary function is not a sufficient condition for normality.

⁴ As Frances – who was in charge of the DSM-IV task force – writes: “DSM-5 has made it easier to diagnose MDD among the bereaved, even in the first weeks after their loss. This was a stubbornly misguided decision in the face of universal opposition from clinicians, professional associations and journals, the press and 100s of 1000s of grieverers from all around the world” (Frances, 2013, 103).

Moreover, the model's validity is challenged by the observation that grounding normality in evolved functions relies on a dubious justification.

2.2 Nesse's Proposal

Nesse aims at providing a robust scientific foundation for psychiatry – and hence at resolving the DSM's limitations – by integrating into psychiatry the principles of evolutionary theory. He claims: “[m]ental health professionals lack a body of knowledge about normal emotional functions.... EP [namely, evolutionary psychology] is the beginning to provide this missing body of knowledge” (Nesse, 2015, 907); and: “[e]volutionary biology... provides, for psychiatry, what physiology provides for the rest of medicine – an understanding of normal functioning as the foundation for understanding pathology” (Nesse, 2017, 769). In other words, Nesse grounds his project in the assumption that our moods and emotions have functions, identified by evolutionary theory, and that we can determine whether an emotional state is pathological by relying on those functions (idem, 769).

A common way of understanding functions in evolutionary theory is the following: the *functions* of a trait or state are “those effects for which [the trait or state] is an adaptation” (Sterenly and Griffith, 1999, 221). An *adaptation* is a trait or state “that exists because natural selection has favored it” (idem, 217) due to its positive effects on fitness, a measure of survival and reproduction. In other words, adaptations are traits or states that were maintained through evolutionary history because individuals capable of carrying or experiencing those were more likely to survive and reproduce, relative to those who were not. The effects on fitness of a trait or state which is an adaptation are its function: a trait is an adaptation because it has a specific function that increases fitness. For instance, jealousy has been claimed to have the function of defending important sexual relations from threats – “with behavior ranging from vigilance to

violence” as Buss holds (Buss, 2013, 156; see Buller, 2005, for a critique of Buss’ position). On this view, jealousy is an adaptation: it was selected for because the individuals capable of jealousy had more descendants and lived longer than those incapable of it.

Conceiving of functions as the effects for which a trait is an adaptation corresponds to the so-called “etiological theory” of functions (Sterenly and Griffith, 1999, 221). This conception is structured around the idea that functions and adaptations are interdependent: an adaptation necessarily has a function and only traits that are adaptations have functions. This view entails that when identifying functions, we are not concerned with a trait’s current ability to increase fitness; according to the etiological theory, *adaptiveness* is as such irrelevant to functions. Adaptiveness is the quality of a trait that is currently adaptive namely, that provides a fitness benefit in the existing environment to the individuals carrying them, relative to those who are not (idem, 217). While etiological theory only cares about adaptations, the difference between adaptiveness and adaptation is crucial. A trait can be an adaptation without being adaptive and *vice versa*. For instance, the ability to read is adaptive but not an adaptation: “[l]iteracy is highly adaptive in most modern human societies, as the disadvantages suffered by dyslexic people testify. But the ability to read is probably a side effect of other, more ancient cognitive abilities” (idem, 218).

But the etiological theory of functions is not the only one used in biology. In fact, it is opposed to two main alternative ways of defining functions: the propensity theory of function (Bigelow and Pargetter, 1987) states that functions should be understood as the “adaptive effects” of a trait (Sterenly and Griffith, 1999, 222), entailing that we should only be concerned with adaptiveness and not adaptations when identifying functions, allowing us to attribute functions to traits which provide a fitness benefit, but that have not been selected. For instance,

bacteria that acquire new traits resulting in antibiotic resistance have a function in the propensity sense, not in the etiological, because that trait is not yet selected for (idem, 222). Finally, functions can be defined without any regard to evolutionary theory: the causal role functions theory (Cummins, 1975) attributes functions to parts of the organism by analyzing their role in the performance of a task, or in the realization of a mechanism.⁵ As Sterenly and Griffith explain, a causal role function can coincide with a trait's etiological function, but not necessarily: the heart's causal role function – to pump blood – corresponds to its etiological function; the heart was selected for to pump blood (idem, 223). However, while “the redness of blood plays an essential causal role function in blushing,... our blood is not red because people who were able to blush had more children than other people” (idem, 223).

Though he is not explicit about it, when Nesse argues that psychiatry should integrate the principles of evolutionary theory to determine the “normal functioning” (Nesse, 2017, 769) of our emotional states, he seems to be referring to the etiological theory of functions. He frequently clearly stresses the link between a state's normality and its evolutionary history, and not merely the ability of a trait to currently increase fitness.⁶ For instance, he claims that “[o]ur knowledge of how natural selection has shaped the body and how it works allows us to be quite confident about what is normal and what is not” (Nesse, Nesse, 2011, 41); knowledge of natural selection indicates us our emotions' “normal functioning” (Nesse, 2017, 769).

Now, how does Nesse describe, more specifically, the normal functioning of our emotions? How do emotional pathologies arise? The basic idea is that an emotion is normal if it

⁵ Propensity theory's functions can in fact be understood as subtypes of the causal role function theory: traits which increase fitness contribute to the overall mechanism of fitness increase of an organism (see Sterenly and Griffith, 223-224).

⁶ Many other claims support this relation, such as: “normal protective responses shaped by natural selection” (Nesse and Stein, 2013, 3), “selection has shaped... mechanisms that often give rise to normal but useless suffering” (Nesse, 2015, 908).

is experienced in accordance with its etiological function. In fact, emotions are carried out by underlying mechanisms: when these mechanisms succeed at expressing the emotions in the right situations – namely, the situations for which the emotion were specifically selected, the situations for which its function provided a fitness benefit – and at the right amount – namely, when the intensity of the emotion is proportional to the situation – the emotion is seen as in accordance with its function, and hence, normal on Nesse’s account. But Nesse argues that these underlying mechanisms are “vulnerable to failure” (Nesse, 2017, 769), or to “dysregulation” (idem, 777), leading to pathological emotional states. More specifically, there are two ways in which these mechanisms can fail. First, the mechanism may fail to trigger the emotion in the right circumstances, namely, either the emotion arises in a kind of situation which is unrelated to the situation for which the emotion was selected or, conversely, the emotion is not triggered in the situation for which it was selected (Nesse, 2009, 23-27). Second, the mechanism can generate an emotional response that is “excessive or deficient” (idem, 27) or “too prolonged for the situation” (Nesse and Stein, 2012, 4).

In other words, there are two ways in which an emotion’s underlying mechanism can fail to carry out an emotion in accordance with its function: the emotion is not *appropriately triggered* – namely, the emotion is provoked in situations that are unrelated to the kinds of situations for which the emotion was selected or the emotion is not provoked in the situation for which it was selected; or the emotion is appropriately triggered but not *proportionate to the situation* – namely, it is deficient or excessive *vis-à-vis* the situation triggering it.⁷ For instance,

⁷ Strictly speaking, the proportionality criterion also not be as such related to whether a state is experienced in accordance with its etiological function: when talking about etiological functions we do not care about how well the trait performs its function – that has more to do with causal role functions and propensity functions – we care about the presence of a trait and whether it is expressed in the right circumstances. But it seems that there is a way to understand the proportionality criterion as proper to the etiological framework: when talking about proportionality, what counts for Nesse is not how well the trait is executed, but rather, if its *evocation is proportional to the situation*, if it is rightly triggered. A more severe situation would trigger a more severe response. In that sense, the

jealousy is normal in situations where there is a perceived threat to a valued relationship – it is aroused to serve its function. But it is abnormal when it occurs in situations where there is no such threat (Nesse and Jackson, 2011, 189) or when the feeling of jealousy is much stronger than what the situation could justify: for instance, if an individual undergoes an overwhelming feeling of jealousy during many weeks just because their partner had a small-talk discussion with another person.

Hence, Nesse stresses the importance of a close examination of the context in which an emotion arises: “[d]etermining if an emotional response is abnormal requires knowing the situation it has been shaped to cope with, and whether or not that situation is present” (Nesse, 2017, 775), and also, as mentioned above, an evaluation of whether the emotion is “disproportionate... to the situation” (Nesse and Jackson, 2011).

Before analyzing how this framework is applied to depression, it is important to note that there is an ambiguity in the way Nesse describes how evolutionary theory should shape our understanding of normality. Nesse does not seem to pay attention to the distinction between adaptiveness and adaptation, and hence to the distinction between etiological and propensity functions. For instance, he mentions: “[i]s a suntan normal or pathological. It depends on... whether the capacity for tanning gives a net advantage or not. In this case, it is assuredly normal” (Nesse, 2001, 41). Here, Nesse defines normality in terms of the adaptiveness of a state; there is no mention of its selected function. Perhaps, the reason why he is dismissive of these distinctions is that he considers that most emotional states’ etiological functions coincide with their propensity function: in virtue of being an adaptation, jealousy is currently adaptive. But not only

proportionality criterion could fit with the etiological framework. The problem is that Nesse does not explicitly explain the criteria he uses; so, the above explanation is my interpretative stance. However, while it can result in a kind of ambiguity of his position, not clarifying what the proportionality criterion amounts to does not seem to generate any kind of inconsistency in his position; that is why I do not focus much on this question.

is that statement often false,⁸ but also, as I will try to show in section 3, failing to explicate these distinctions leads to inconsistent claims regarding the normality of emotional states that are not adaptive in the existing environment, but normally triggered to perform their etiological function.

How is Nesse's understanding of the normal functioning of our emotions applied to the demarcation problem of depression? According to Nesse, depression mainly arises from the failure of the mechanism underlying low mood, an emotional response shaped by natural selection (Nesse, 2017, 777). Indeed, low mood, described as "the cluster of symptoms usually associated with depression – depressed mood, anhedonia, crying, self-reproach, fatigue, pessimism, psychomotor retardation, somatic disturbances, and shifts in cognitive style" (Keller and Nesse, 2005, 28) was selected for because experiencing it allowed individuals to withdraw from unattainable goals (Nesse, 2009, 23): an unattainable goal would often be harmful to fitness, given that an individual would be likely to spend cognitive and metabolic resources needed for survival and reproduction without a pay-off. Low mood is a protective mechanism that has a de-motivational function: disengaging from an unattainable goal could ultimately lead to a relocation of resources to more accessible and fruitful goals. Nesse explains:

Low mood is aroused when efforts to reach a goal are proving fruitless. That situation has recurred millions of times for individuals in the course of human evolution. While simply persisting with a positive attitude seems intuitively attractive, that is not the best strategy when foraging in the winter, trapped in the pursuit of an uninterested partner, or trying for the fifth year in a row to get into medical school. In such situations, it is best to pause, conserve energy, consider other strategies, and, if no route appears viable after all options have been explored, give up (Nesse, 2017, 776).

This evolutionary explanation of low mood has, in fact, received significant empirical support. For instance, studies have claimed that individuals who are more likely to quickly

⁸ Traits can acquire new functions through the process of exaptation: for instance, "[m]ammal ear bones are converted jaw bones; they are exapted for hearing" (Sternerly and Griffith, 1999, 219), or simply acquire new functions they never had, as in the example of the bacteria becoming resistant to anti-biotics mentioned above.

disengage from a goal are less inclined to experience frequent states of low mood (Wrosch et al., 2003).⁹

When is low mood more specifically abnormal? Low mood is abnormal when it is not in accordance with its etiological function. Hence it is abnormal either when it is not appropriately triggered – namely, when provoked in situations unrelated to those for which low mood was selected;¹⁰ that is, in situations in which one is not trapped in the pursuit of an unreachable goal – or when appropriately triggered but disproportionate to the circumstances triggering it. In other words, low mood is normal when it is experienced in situations in which one is pursuing an unreachable goal. If it arises in a situation where one is not pursuing an unreachable goal, it is likely to be pathological. It is also pathological in cases where the individual is pursuing an unreachable goal, but their emotional response is excessive, or deficient: suppose, for instance, that a lawyer, having won over fifty cases in their life, is experiencing intense loss of motivation, pessimism, and low energy, after trying in vain to win one particular case which has low stakes both for them and the person they are defending. It seems that we could consider this reaction as excessive to the situation, and hence, pathological on Nesse's terms.

Finally, Nesse adds one additional criterion for pathological low mood: “[i]f pursuit of an unreachable goal continues, ordinary low mood can escalate into severe depression” (Nesse, 2009, 24). If Y has been trying for years unsuccessfully to move to another country, it is likely that their original low mood will become pathological. This final criterion seems to have been

⁹ There are reasons to question whether the totality of the evidence Nesse mentions can actually function as genuine support to his position. For instance, Wrosch et al's (2003) study seems to suggest that low mood arises as the result of a failure to disengage from a goal, which would entail that it is not what motivates the individual to withdraw from a goal, as Nesse claims. Unfortunately, while examining the empirical evidence for Nesse's explanation is crucial, it cannot be treated here.

¹⁰ In accordance with what was said above, there is also a pathology of low mood when low mood is not triggered in the situations that are supposed to trigger it (Nesse, 2009, 23). I will not, however, mention it longer here, because while it could be a pathology of low mood for Nesse, it cannot be abnormal low mood as such, namely, depression, because low mood is absent.

added by Nesse to include empirical data on the kinds of situations leading to depression, mainly gathered by Klinger (1975) who seemed to be the first to observe a correlation between low mood and goal pursuit and to suggest that low mood could be an adaptation¹¹ (Nesse, 2009, 23). However, as I will show in the next section, this final criterion seems at odds with Nesse's core ideas on the relation of normality to etiological functions.

3 CAN EVOLUTIONARY THEORY GENERATE A VALID MODEL TO DEMARCATHE THE NORMAL FROM THE PATHOLOGICAL?

Nesse's model does seem to fill in the DSM's gaps: first, it reconciles the demarcation criterion with the common idea that some emotional states are normal reactions to the circumstances, and hence, that the context in which the symptoms occur should be considered during diagnosis. While the DSM-5 would encourage diagnosing with major depressive disorder a mother who is helplessly trying to help their child recover from drug addiction, after two weeks of experiencing depressive symptoms, Nesse's model would first encourage her to understand that her state is a normal reaction to this difficult situation, because our bodies have been shaped by evolution to react in such a way, in light of the kinds of situation she is facing. It is important to note that Nesse's diagnosis does not mean that the mother should not receive treatment: "people deserve relief from depression whether or not it arises from a[n] abnormality. Physicians in the rest of medicine use medications to relieve pain, whether it is aroused by an appropriate stimulus, or by some abnormality in the pain system" (Nesse, 2009b, 472).

¹¹ Nesse writes: "A seminal article by Klinger (1975) initiated modern work on mood as an adaptation that regulates goal pursuit. He noted that rapid progress towards a goal arouses high mood that motivates continued effort and risk-taking. When efforts to reach a goal are failing, low mood motivates pulling back to conserve resources and reconsider options. If the individual persists in an unreachable goal, ordinary negative affect can escalate into pathological depression" (Nesse, 2009, 23).

Secondly, it seems to provide an objective theoretical demarcation criterion: etiological functions are biological facts. With few exceptions – due to genetic or phenotypic variations – all humans carry the same adaptations with the same functions: insofar as they are adaptations, they have been fixed in the population. According to Nesse, identifying them does not depend on practitioners' individual preferences or cultural and socio-economic determinations because all functions are triggered and expressed in the same circumstances across all members of a species. For instance, as Nesse claims, “we all have salivary glands that make secretions” in the same conditions, namely “when we anticipate or eat food”: producing secretions in those situations is the function of salivary glands; this is “objective and depends on no social input” (Nesse, 2001, 44).

However, the application of Nesse's evolutionary model for the demarcation of the normal from the pathological may face some impediments in the practical setting of the diagnosis of depression. A clinician applying Nesse's model to judge whether their patient is suffering from pathological depression would need to correctly identify whether there is a goal the individual has been in vain trying to achieve – whether the triggering circumstances are there – and also to assess whether the emotional response the patient undergoes is proportionate to the situation. The difficulty is that both these steps heavily rely on the practitioner's subjective judgment. First, not only many situations could be interpreted as the pursuit of a goal, but what counts as an “unattainable” goal does not seem evident: how do we know after how many failed attempts at entering medical school does the goal of becoming a medical practitioner become unattainable? Second, to assess if an emotion is proportional to the situation presupposes judging the degree of severity of a situation, the degree of intensity of an emotion and evaluating whether the two coincide. The practitioner's own experiences are very likely to interfere with this

complex judgment. For instance, one clinician may consider that one's failed attempts to win a scholarship might not be a reason for an important mood drop, where the other might disagree.

If determining whether a patient is suffering from a pathology ultimately relies on the practitioner's subjective judgment on whether the symptoms are proportional to the situation and whether the situation does correspond to the pursuit of an unreachable goal, then diagnoses across practitioners are very likely to be inconsistent. This inconsistency across practitioners is precisely what the DSM aimed at resolving (Nesse and Jackson, 2011) – in this sense, Nesse's model could be seen as a throwback. But Nesse recognizes this difficulty and supports the idea that further research could overcome the problem of subjectivity. In particular, he claims: “subjective judgment [about the severity of the life situation] is unavoidable.... Change will eventually come as researchers discover that their findings become stronger when they differentiate subpopulations according to how disproportionate symptoms are to the situation” (Nesse and Jackson, 2011, 175).

Hence, while the practical utility of Nesse's model is a pressing question – in fact, the relevance of his position is put in doubt if it cannot technically lead to objective diagnosis – it is not essentially what concerns us here. What I will try to show in the next section is that there is a theoretical inconsistency in Nesse's model which reveals that low mood's functioning in conformity with its function cannot be a sufficient condition for normality, thus challenging the foundation of his project.

3.1 Is Low Mood's Conformity with its Etiological Function a Sufficient Condition for its Normality?

Let us briefly review the cases in which low mood is either normal or pathological according to Nesse. There are two main general conditions determining the normality or abnormality of low mood.

1) Failure (or dysregulation) of low mood mechanism condition. If the normal mechanism of low mood fails, then low mood can be pathological. There are two ways in which the mechanism can fail: 1a) and 1b).

1a) Low mood is inappropriately triggered. If low mood is present in a situation which is unrelated to the situation for which its function was selected, it is pathological. Conversely, low mood is normal when it is present in a situation for which its function has been selected and when it is absent in a situation unrelated to its function.

Condition 1a) needs to be further specified by condition 1b): if low mood is present in a situation with which it has been shaped to cope, its normality is further determined relatively to its proportionality to the situation.

1b) Low mood is disproportionate to the situation condition. If low mood is disproportionate to the situation, it is pathological. Conversely, if low mood is proportionate to the situation, it is normal.

2) Persistence condition. If X continues to persist in their unreachable goal, low mood can turn into pathological depression.

Following Nesse's example, suppose Zehiri has been "trying for the fifth year in a row to get into medical school" (Nesse, 2017, 776). For the past three years, he has been experiencing low mood, decrease in motivation, smaller appetite, and lack of interest in the activities that usually pleased him. The symptoms have been exacerbating over the past two years. Is Zehiri suffering from pathological depression?

According to condition 1), his emotional state is not pathological: he experiences low mood because he is trapped in a goal he cannot achieve. His low mood is appropriately triggered because it is present in a situation for which its function has been shaped to cope with. We cannot also claim that his reaction is disproportionate to the situation: the longer the situation lasts, the more his symptoms are intensifying. However, according to condition 2), it is likely that Zehiri's low mood is pathological. Despite him consistently feeling low, he has not disengaged from his goal. His low mood has "escalated into clinical depression" (Nesse and Jackson, 2011, 776). But Zehiri's state cannot be both non-pathological and pathological. And while on the first condition Zehiri's state is accepted as normal, it seems that cases such as his are the reason there is a second main condition for pathology. As mentioned earlier, Nesse adds this condition as a result of important empirical studies on depression, which show a correlation between goal persistence and low mood levels (Klinger, 1975, Nesse, 2009). I will therefore consider that Zehiri's case is, in fact, pathological.

This example helps to point to an inconsistency in Nesse's model: according to the failure condition, the normality or pathology of a state is determined relative to the state's conformity to its etiological function. Essentially, a low mood state is normal when it is appropriately expressed in response to a situation with which it has been shaped to cope. But this condition is in contradiction with the claim that an individual can experience a pathological state when they

continue to pursue the unreachable goal despite their symptoms intensifying. In the latter case, low mood is still appropriately triggered by, and proportional to the situation. There seems to be nothing in the mechanism failure condition that can explain why this case is pathological.

What this inconsistency seems to reveal is that, even by Nesse's own understanding of depression, there is more to the pathological than mere deviance from or unconformity with etiological functions. To explain why low mood is pathological if a person continues to pursue an unreachable goal, Nesse cannot rely on the assumption that pathology is defined as the disruption of a normal functioning, a functioning in accordance with the mechanisms shaped by natural selection, because low mood is still appropriately triggered and proportionally experienced in that situation. In other words, there seems to be an explanatory gap in Nesse's model.

In response, Nesse could argue that this explanatory gap does not challenge his overall demarcation project. In fact, Nesse never claimed that failure of the normal mechanism of low mood is the only cause of pathological depression. As he mentions, pathological depression can "aris[e] from mechanisms unrelated to those that normally regulate mood" (Nesse, 2009, 27), such as "brain or cognitive abnormalities" (Nesse, 2015, 910). Since there are other causes of depression besides the failure of the mechanism expressing low mood in accordance with its etiological function, Zehiri's case could simply be an instance of a different kind of process causing depression.

But admitting this conclusion amounts to significantly reducing Nesse's model's scope: as I will try to show, Nesse does not have the theoretical tools to demarcate the normal from the pathological in cases such as Zehiri's; in fact, the demarcation of the normal from the pathological in cases where "[c]ontinued pursuit of an unreachable goal can escalate into clinical

depression” (Nesse, 2017, 776), can only be done, it seems, with an approach to diagnosis solely based on the presence and intensity of the symptoms, as the DSM pursues. To account for the reason why Zehiri’s case is pathological, Nesse would need to identify a cause that turns low mood into depression – a cause which cannot be the dysregulation of the low mood mechanism, given that the latter is still functioning in accordance with low mood’s function: it is appropriately triggered and proportional to the situation. This cause might just be the failure to disengage from the unreachable goal. But Zehiri has not disengaged from this goal for the past five years, so at what moment exactly did the failure to disengage cause a pathology? Given that Zehiri’s situation has not changed, it seems that only available tool to identify the moment in which normal low mood turns into a pathology is the evaluation of the symptom’s evolution. In other words, to demarcate normal low mood from pathological in Zehiri’s case, we would be relying on the presence and intensity of the symptoms. But this is precisely how the DSM operates and what Nesse wants to avoid (Nesse and Stein, 2012).

The solution would be to find a theoretical explanation of the mechanism which converts normal low mood to pathological depression in cases such as Zehiri’s. Until such an explanation is given, it is difficult to see how diagnosing depression by following Nesse’s model could be different from an atheoretical, symptom-based approach in those cases.

However, there seems to be a deeper issue with Nesse’s project. Nesse’s project of providing a scientific foundation for psychiatric diagnosis and classification essentially relies on the assumption that a state is normal if its presence is in conformity with its etiological function: “[p]sychiatric nosology is constrained by the lack [of] a functional understanding of normal behavior” (Nesse and Stein, 2012, 6). Defining normality in terms of etiological functions is at

the core of Nesse's project of a genuinely scientific psychiatry and at the root of his demarcation model.

But the crucial assumption of defining normality in terms of etiological functions is challenged. Cases such as Zehiri's, where "[c]ontinued pursuit of an unreachable goal can escalate into clinical depression" (Nesse, 2017, 776), show that low mood can be appropriately provoked, but abnormal. Hence, appropriateness of a state to the situation in accordance with its etiological function is not a sufficient condition to determine normality. Nesse could reply that we need not hold that a state's conformity with its etiological function is a sufficient condition for normality; it is enough to say that it is a necessary condition. Indeed, we could say that what matters is that we cannot have normality without proper functioning, and that proper functioning does not need to guarantee normality. But admitting this undermines the foundation of Nesse's project.

What matters for Nesse is not to define normality for itself, but to define normality in order to understand pathology. If we only aimed at understanding normality, perhaps, a necessary condition for it would suffice. We would be able to formulate something like the general law "all normal states have proper functioning". But that is not what Nesse is looking for. Indeed, at the core of his project is the idea that the understanding of normal states is the principle from which the way to demarcate the normal from the pathological is derived. Nesse writes:

A deep understanding of what is abnormal requires a richly detailed understanding of what is normal... we require nothing less than a complete knowledge of what the body is for, how it works, and, especially, how it came to have its current form. If we had this knowledge in hand, then we could define abnormality with reference to deviations from normality, not needing to resort to either statistical or value laden information (Nesse, 2001, 38).

If proper functioning cannot guarantee normality – if it is not a sufficient condition for normality – as Zehiri’s case shows, we lose the essential criterion by which we can demarcate normality from pathology, the principle underlying Nesse’s project: states can be pathological, even if they function in conformity with their etiological functions.

But Nesse could also argue that Zehiri’s case is not pathological. Zehiri might experience discomfort, but he is not suffering from a pathology. If Nesse admits that – at the cost of contradicting the research, mainly Klinger’s work (1975), he himself considers fundamental on depression and its etiological function (Nesse, 2009, 23) – then the inconsistency is resolved. Thus, supposing that there is nothing in Nesse’s writings that would entail that functioning in conformity with etiological functions is not sufficient for normality, are we still justified in grounding the definition of normality in etiological functions? In the next section, I will try to show that the justification for this claim is dubious.

3.2 Can we Ground Normality in Etiological Functions?

Nesse’s core idea is that we can understand the difference between pathology and normality by relying on our state’s etiological functions: the normality of a state is defined as what is expressed in conformity with its etiological function (Nesse, 2017). This core claim has two implications. First, it entails that normality is restricted only to those states that have a function: consequently, states that do not have an etiological function – namely, which are not adaptations – are not prone to be normal. Second, it seems to presuppose that the normality of a state in the current environment is assessed based on our understanding of how that state functioned in the past, at the time it was selected for. Namely, it presupposes that our emotion’s etiological functions coincide with their causal role functions. In this section, I will examine these two implications. I will try to show that, while the validity of the first implication can only

be settled with empirical studies, the second implication is straightforwardly problematic: there is a discrepancy between our emotions' etiological function and their causal role function. Given this discrepancy, it is not reasonable to determine a state's normality relative to its etiological function.

3.2.1. Are Non-selected States Pathologies?

Can emotional states be normal even if they do not have an etiological function? Nesse's understanding of psychiatric pathology – but also of medical pathology, in general – relies on the view that most pathologies stem from mechanisms shaped by natural selection. He explicitly states: “[t]he first task in an evolutionary analysis of a medical condition is to determine if it arises directly from a bodily defect or if it is an adaptive response to a more fundamental problem” (Nesse, 2009, 20). Diseases arise either from bodily defects, or from the failure of selected mechanisms. But this claim is a false dichotomy: it excludes the possibility that diseases originate from mechanisms that are not shaped by natural selection, but that are the by-product of other traits, traits accidentally maintained throughout evolution as a product of genetic drift or traits shaped by developmental factors (Lewontin and Gould, 1979): for instance, many support the idea that various monogenic diseases, such as Tay Sachs disease – a disorder damaging nerve cells in the brain or the spinal cord – or cystic fibrosis – a disorder mostly affecting the lungs – do not originate from the failure of a selected mechanism but by the impact of random genetic drift (Valles, 2010).

Thinking that disorders arise essentially from mechanisms shaped by natural selection relies on the idea that most states are shaped by natural selection. In other words, Nesse's writings seem to reflect a commitment to adaptationism, the view distinguished, among other things, by its “unwillingness to consider alternatives to adaptive stories” (Lewontin and Gould,

1979, 581). Not all human traits have been shaped by natural selection. It is likely that Nesse has a biased view of natural selection's scope, which is likely to result in a biased view of normality: if most of our traits and states have been shaped by natural selection, it can be reasonable to ground normality in selected functions. But if it is not the case, if a lot of our states are the result of other evolutionary processes such as random genetic drift, then grounding normality in natural selection would be problematic: it would mean excluding an important number of states from normality.

In fact, there are traits that are ordinarily considered normal, for which it is difficult to maintain that they were shaped by natural selection. For instance, Nesse explains:

[I]f lack of sexual interest in members of the opposite sex is shown to result from autoimmune damage to a particular part of the brain, then it is abnormal. If, however, it is shown to be a facultative adaptation that is aroused only in certain circumstances where it contributes to kin survival and reproduction, then it would be an adaptation. Although Wilson suggested this second possibility, I know of no evidence for it (Nesse, 2001, 41; internal citations omitted).

In other words, if we cannot show that lack of attraction to the opposite sex has a function, Nesse's model pushes us to believe that it is abnormal; homosexuality could be used as a counterexample to his position. Now, there are newer evolutionary explanations of homosexuality based on natural selection, according to which same-sex attraction was selected for because it "strengthened social bonds and reduced conflicts between women in polygynous marriages" (Luoto, 2020, 2241). It is hence possible that homosexuality cannot constitute a counterexample to Nesse's model. Indeed, identifying such counterexamples to refute Nesse's position runs into important difficulties: mainly, the debate on whether many of our traits are adaptations is empirically open (Pradeu, 2011, 387). Given that this debate can only be resolved through further empirical studies, I will not focus longer on it.

The question that interests me is whether we are justified in determining our traits and states' normality based on the way our bodies have been shaped by evolutionary forces. If not, then it is not easy to see why relying on evolutionary principles at all can be a good candidate for the definition of normality. I will argue that there is a significant gap between how our emotions have been shaped by evolutionary forces, and how our bodies actually operate, in the existing environment. In other words, I will try to show that there our emotions' etiological functions are often at odds with their causal role functions, a discrepancy which is especially noticeable in the case of low mood. If that is right, there does not seem to be any obvious legitimate justification of why we would determine the normality of our emotional states according to standards derived from an understanding of their etiological function.

3.2.1. A Discrepancy between Etiological and Causal Role Functions

Nesse indicates that there is a current increase in depression's prevalence, which can be explained by the fact that our environment is substantially different from the ancestral environment in which low mood was selected for: “[d]epressogenic situations may be especially common in modern life because goals are far larger and longer in duration than those the regulation mechanism was shaped for” (Nesse, 2009, 25). Nesse appeals to the concept of an evolutionary “mismatch” (Nesse, 2015, 904) to account for the modern increase in pathological cases of depression. An evolutionary mismatch corresponds to a situation in which: “adaptations that contributed to survival and reproduction in previous environments become relatively maladaptive in a changed environment” (Lloyd, Wilson and Sober, 2011, 3).

The presence of such a mismatch precisely shows two important discrepancies. First, low mood's etiological function – the motivation to disengage from unreachable goals – does not coincide with its propensity function: presumably, low mood does not have a propensity function

at all. The mismatch hypothesis entails that low mood is no longer properly adaptive in our environment; it is no longer a determining force for us to disengage from our goals as it were for our ancestors. Second, low mood's etiological function does not coincide with its causal role function: if low mood is not what motivates us to actually disengage from our goals, then, it simply does not operate as it did in the past, at the time it was selected for. The role that low mood plays in our current bodies and environments does not map onto its etiological function, the function that it had when it was selected for. Given those discrepancies, it seems likely that we do not have any legitimate reasons to determine low mood's normal expressions as those that are in conformity with its etiological function: low mood simply does not function as it did in the past. In fact, physiology, which Nesse takes as the standard towards which psychiatry should tend, acknowledges that there can be a discrepancy between etiological functions and the way in which bodily mechanisms currently operate. Physiology only cares about causal role functions, not etiological functions. Indeed, as Sterenly and Griffith explain:

[P]hysiologists... are not typically concerned with evolutionary history. They are interested in the activities an organism can perform: flying, digesting food, detecting viruses in its tissues, and the like. They explain how organisms perform these activities by functional analysis – by breaking down the overall task into parts that are performed by different parts of the organism (Sterenly, Griffith, 1999, 223).

If psychiatry were truly to mirror physiology as Nesse wishes, then it would rely solely on their causal role function, not their etiological function; psychiatry would need to analyze the role played by our emotions in our organisms, and determine their normality relatively to that, regardless of its past or present adaptive value.

One can, however, remain unconvinced by this line of thinking and argue the following. While low mood's etiological function might not coincide with its propensity function, it is not necessarily at odds with its causal role function. In fact, there is no discrepancy between the way

our low mood mechanism functioned in the past and how it functions now: our low mood still functions in the same way, insofar as it is triggered by the same situations – those for which it was selected – and it still, while much less successfully, motivates us to disengage from unreachable goals. It is still triggered and expressed according to its etiological function. Low mood's function has not changed. It is only the environment that has changed: the environment is to blame for its being less effective. Hence, the normality of our low mood can still be determined according to its etiological function. As Nesse explains, the change in the environment has simply the unfortunate consequence that low mood states are more likely to become pathological. If our ancestors were living in our modern environment, they would experience low mood in the same way we do, and they would also be more prone to experience pathological depression.

But this response faces two shortcomings. First, it seems that the pathological nature of depression resulting from the mismatch between our modern environment and the selected mechanism of low mood cannot be explained by appealing to Nesse's core criterion for normality. Admitting that cases of low mood can be pathological because of this mismatch poses the threats to the logical consistency of Nesse's position described in section 3, namely that states can be both pathological and non-pathological at the same time. Indeed, Nesse explains that the main reason why our modern environment is more conducive to pathological cases is that modern's life goals are "far larger and longer in duration than those the regulation mechanism was shaped for" (Nesse, 2009, 25). Hence, an individual is more likely to be trapped in an unattainable goal, thus increasing their chances of becoming pathologically depressed. However, in that case, low mood is still appropriately triggered and proportionate to the situation: in this sense, low mood is still normal. In other words, those instances are normal

according to Nesse's first condition for normality – the failure condition – but abnormal according to how Nesse interprets the evolutionary mismatch – and also abnormal according to Nesse's second condition, the continued persistence condition, mentioned in section 3. Here, the distinction between adaptive value and adaptation and etiological functions and propensity functions becomes crucial. Failing to explicate those distinctions, Nesse arrives at conflicting observations: he claims that low mood is both abnormal when no longer adaptive, and normal in those same instances, because triggered and expressed in conformity with its etiological function. In other words, Nesse cannot have a straightforward, coherent answer to the question posed in the first section – namely, are we currently undergoing a depression epidemic; is there a substantial increase in pathological cases in the world? – because his criteria for normality imply that, while there is an increase in depression cases, the increase is at the same time normal and abnormal.

Second, grounding normality in evolved functions presupposes a kind of fixity of our evolved states. Assuming this fixity, especially regarding our emotional states, seems unwarranted: our bodies evolve in response to the environment; if the environment is changing, there is no reason to assume that our emotions are not changing as well. It seems difficult to assume that our emotional states have not changed at all in the light of environmental changes. It is, therefore, not satisfying to claim that low mood's etiological function maps onto its causal role function.

4 A POTENTIAL SOLUTION: REFORMULATING THE NORMAL-ABNORMAL DISTINCTION IN PSYCHIATRY?

The previous section aimed at pointing to the limitations of using evolutionary principles in psychiatry as tools to distinguish normal cases of low mood from abnormal depression, as Nesse suggests (Nesse, 2009, 2012, 2017): low mood's expression in conformity to its etiological function cannot guarantee its normality. Moreover, grounding normality in etiological functions presupposes a kind of fixity of our emotional states which seems unwarranted in the light of a continually changing environment. In this final section, the goal is to suggest a possible alternative to Nesse's position for the demarcation problem of depression.

4.1 Abandoning the Functionalist Framework?

If it is not reasonable to ground normality in etiological functions because the way they currently operate is distinct from how they did in the past, then we can perhaps rely on causal role functions, in a way that faithfully mirrors the way physiology works, to demarcate normality from abnormality. We could identify each of our basic emotions' causal role function through psychological studies, obtain a body of knowledge on the role each emotion performs in the body, and rely on those to determine their normality: emotions which do not conform to their psychological causal role function would be abnormal. However, even if we could in theory obtain such a body of knowledge, – it would remain doubtful whether that could provide us with a valid demarcation criterion for emotional disorders, because, as I will try to show, an emotion's conformity with its causal role function cannot guarantee normality either.

Suppose that we know that disgust has the causal role function of motivating us to avoid toxic objects or situations. In a vein similar to Nesse's, disgust would be abnormal when not expressed in conformity with its causal role function. But suppose that the reason why a person

fails to experience disgust is because they have adopted a principle according to which all forms of disgust should be eliminated because disgust is thought of as a useless and harmful emotion. Belief in this principle and years of training to apply it, have resulted in the individual never experiencing disgust. They have learned not to feel disgust in toxic situations that should elicit disgust according to its causal role function and have avoided ever getting intoxicated without experiencing disgust. There is no reason to assume that their case is pathological: they have rationally imposed on themselves a principle, which directly acts on the way their emotions are expressed.

In *The Normal and the Pathological*, Canguilhem seems to point to the idea that normality cannot be defined solely in terms of conformity to functions. For instance, he mentions Hindu yogis who, through meditation, gained the ability to modify how their physiological functions operate:

To obtain a change in pulse rhythm from 50 to 150, an apnea [absence of respiration] of 15 minutes, an almost total suppression of cardiac contraction, certainly amounts to breaking physiological norms. Unless one chooses to consider such results pathological. But this is clearly impossible (Canguilhem, 1991 [1966], 165-166).

The idea here is that humans have the ability to act on the way their emotions are expressed and even on the circumstances which trigger them, and, hence, to digress from their fixed functions. In fact, for Canguilhem, a healthy subject is one which precisely digresses from rigid norms, not one which acts in total conformity with those: “being healthy and being normal are not altogether equivalent.... What characterizes health is the possibility of transcending the norm, which defines the momentary normal, the possibility of tolerating infractions of the habitual norm and instituting new norms in new situations” (ibid, 196-197). Canguilhem’s position points to the functionalist approach’s important limitations for demarcating the normal

from the pathological, whether the functions it appeals to are evolutionary or causal role functions.

A possible solution would be to abandon the functionalist framework altogether: the demarcation problem in depression diagnosis – and maybe overall in psychiatric diagnosis – cannot be solved by appealing to functions, whether those are evolutionary or not. Moreover, it seems that there is a way to avoid the functionalist drawbacks, while remaining faithful to one of Nesse's core requirements for the differentiation of normal from abnormal states: the importance of considering the context in which emotions arise when diagnosing mental disorders. In particular, the contextual approach claims that:

Depression symptoms that have some sufficient justification... and, crucially, that are a proportional response to that justifying trigger, are not indicative of illness. Conversely, those symptoms which do not have sufficient justification... or which are a disproportional response to a justifying trigger, are indicative of illness (Tully, 2018, 116).

Nesse's functionalist approach could be understood as a species of the contextual approach. The difference between the two is that Nesse appeals to low mood's etiological function to pinpoint a specific type of situation which would normally trigger low mood: the pursuit of an unreachable goal. However, since pursuing a goal is an extremely broad situation – a wide range of situations can be read as the pursuit of a goal – in a clinical setting, the difference between Nesse's take on the demarcation problem and what is suggested by the contextual approach is often likely to be unnoticeable. For example, if a student is experiencing heavy low mood, lack of motivation, disinterest in the activities that usually please them for a few weeks before a very difficult exam period, both would consider their case normal, because proportionate to the situation. Ultimately, both Nesse and a defender of the more general contextual approach would determine whether the emotional state is pathological based on

whether the depression-like symptoms are proportionate to the situation, in the sense that the intensity of patient's symptoms corresponds to the degree of severity of the situation they are in. Moreover, it seems that considering the environment in assessing the normality of a person's emotional states, is not a new idea, but common practice in psychotherapy.

4.2 Normality-Abnormality: an Irrelevant Distinction?

However, it seems that the contextual approach has profound limitations. There are psychological states which are justified responses to the environment, but which we consider disorders. For instance, there is a wide consensus that multiple personality disorder – characterized by “simultaneous manifestation of multiple alternative personalities in one human body” (Cudzik, Soroka and Olajossy, 2019, 1) – emerges as a defense mechanism in response to a traumatic situation:

[M]ultiple personality disorder can be a broad variant of the child's defense mechanisms against extreme, traumatic events from childhood.... Abused children create other representations of the Self to be able to rid themselves of suffering, a process that is necessary for them to survive and further develop mentally and physically (idem, 1).

It seems that there is no space, in the contextual approach, for pathologies which do not arise from any internal dysregulation of functional mechanisms but that are directly caused by the context. In other words, in some sense, the only states that are legitimately called “pathologies” are those in which the pathology is generated from the individual itself, and not caused by external factors. If normality is always defined as what is proportionate to the situation, no matter how damaging the situation may be on an individual, the emotions generated from that damage could always be considered normal, being proportionate responses to the situation. Such an approach cannot explain the pathological nature of cases such as multiple personality disorder.

The functionalist approach and the broader contextual approach often seem to lead to conclusions that are either internally inconsistent or at odds with common scientific treatments of many disorders. Nesse's position entails that Zehiri's case is at the same time normal – because in adequacy with its etiological function, Nesse's core criterion for normality – and abnormal, for reasons unrelated and incompatible with Nesse's core criterion: empirical observations establish a correlation between persistence in unreachable goals and pathological depression. Similarly, the contextual approach entails that conditions such as multiple personality disorder, stemming from a defense mechanism, cannot be pathological. While I could not find any proponent of the contextual approach explicitly discussing the pathological nature of multiple personality disorder, I think most would agree that it is pathological, based on the severity, the high-level impairment it inflicts, and rarity of the condition, despite its functioning as a protective mechanism.

Perhaps, a straightforward way to avoid this inconsistency is to establish that the distinction that matters in psychiatric diagnosis is not “normality” *versus* “abnormality”, but “that which requires intervention” *versus* “that which does not”: for if divergence from normality does not straightforwardly indicate a disorder, the abnormal-normal distinction might just be an abstract theoretical construct which cannot inform clinical practice.

4.3 An Intervention-based Approach

To summarize what was said above, it seems that there are two important problems with using a normality-abnormality distinction as the determining factor in a clinical psychiatric framework. First, when determining abnormality in relation to fixed functions, we undermine our capacity of overcoming fixed norms and instituting our own, a sign of health, as Canguilhem suggests. A healthy individual is not one that systematically conforms to norms. Second, many

cases, such as Zehiri's, are ambiguously deemed both normal – because they conform to a function or because they are justified by the context – and abnormal, for empirical reasons, mostly related to the severity of the symptoms. But a final consideration should be added here: as Nesse acknowledges, the normality-abnormality distinction in psychiatry does not guide clinical treatment (Nesse, 2009b). Both normal and abnormal cases of low mood, for instance, might need to receive treatment. I will try to show that by switching from a framework in which the normality-abnormality distinction is central, to one in which an intervention criterion is central in psychiatry, we can hope to avoid these issues.

Indeed, the aim of psychiatric practice is, similarly to the one in medicine, I think, to provide treatment to individuals in need, not to determine whether their state is normal or abnormal as such: the normal-abnormal distinction is relevant only if it can be used to inform clinical practice. Suppose a strange skin rash appears on X's back. If X visits a medical practitioner for their skin rash, they would do so with the aim of knowing whether that skin rash is a threat to their health. If they are worried about whether their rash is normal, they would be so, it seems, in virtue of wanting to determine whether they would need to do something about the rash to avoid any health problems it could cause. In a similar vein, think of an individual visiting a psychiatrist because of an overwhelming feeling of anxiety: they go to the practitioner to find relief, not to determine whether what they are feeling is normal – though it is true, as Nesse claimed (Nesse, 2009b), that if they are told that what they feel is normal, it might help them feel better. The core of medical practice is therapeutic: it is not about determining whether states are normal or abnormal, it is about finding ways to treat states that need treatment. Determining the normality and abnormality of states – if at all possible, in an objective way – might be a goal for other disciplines, such as physiology, but not for medicine.

For those reasons, it seems that when a practitioner examines a patient in psychiatry, they should not look to determine whether their patient's state is abnormal; they should be evaluating whether the patient finds themselves in need of a kind of intervention – whether it is a drug-based treatment, a form of therapy or counseling – and if needed, they should determine with the patient what kind of intervention they would need. The basic idea is that an individual is in need of intervention in psychiatry when they are stuck in an emotional situation – a set of various emotions – which they seem unable to manage or overcome. They find themselves in a condition of resisting emotional fixity. In some sense, the individual loses themselves. For instance, if Z has been consistently experiencing for some weeks depression-like symptoms – lack of appetite, disinterest in activities that usually pleased them, heavy low mood – and they feel that they cannot get rid of those feelings, despite wanting to, Z needs some form of intervention. Whether that emotional condition is directly caused by the situation the individual finds themselves in or not, is not what matters most: the individual might be justified in feeling that way by the context or they might not. It is only in virtue of the fact that they are experiencing a resting emotional fixity that they need intervention. Contrary to what Nesse believes, evaluating the symptoms in light of the context is not the most essential in psychiatric diagnosis: ultimately, the intervention-based position seems closer to the DSM's recommendations to focus on the symptoms and essentially disregard the context.

The difficulty of this intervention-based account is to determine what experiencing a resisting emotional fixity amounts to and how to identify it. While spelling out what it means would require more work, the basic idea is that it refers to a constant and long-lasting conflict between how the person wants to feel and the emotions they undergo. This approach is in fact indirectly related to Canguilhem's understanding of health. When explaining Canguilhem's

conception of health, Magree writes: “[h]ealth as such is a creative, propulsive, and dynamic state. It is fundamentally opposed to the adoption of a way of being that is fixed or static” (Magree, 2002, 201). The individual needs intervention when they cannot overcome an emotional fixity, an unhealthy state.

How would this framework apply to Zehiri’s case, the student who has been trying in vain to enter medical school for five years in a row? Zehiri does seem to be in a condition of resisting emotional fixity: for the past few years, he has been experiencing depression-like symptoms and has not found a way to overcome them. For those reasons, Zehiri would need a kind of intervention. Furthermore, what would the intervention-based account say about the opening question related to the worldwide increase in depression cases? By depression we now mean a state that requires intervention, described by an emotional situation characterized by a set of depression-like symptoms – such as low mood, lack of appetite, loss of motivation, anhedonia, suicidal ideation, etc. If we agree with Nesse’s claim that low mood is in fact related to continued pursuit of unreachable goals – we can accept that claim without adhering to his conception of normality – and with the observation that goals in modern life are longer and larger, the intervention-based approach would conclude that the increase in depression cases is, at least partly, not artificial. It corresponds to a real increase in cases that require intervention. If humans find themselves more often trapped in situations generating depression-like symptoms, the risk of feeling blocked in their emotions is likely to increase. The advantage of the interventionist-based approach is that it allows us to make a judgment about these cases – both cases similar to Zehiri’s and cases related to the modern conditions of our environment – without arriving at inconsistent conclusions: those cases might be justified in light of the context, but in need of a kind of intervention.

There are many objections that could be directed against this account; but this section is just a sketch of the interventionist approach, which would need to be carefully developed in order to be considered. Hence, I will only briefly discuss one that strikes me as the most important. A first immediate objection of this account would be that it heavily relies on the individual's own awareness of their emotional states, a kind of awareness that many patients might not have: it requires that one can know, or feel, to a minor extent, at least, that one is blocked in a resisting emotional fixity. But the following response can be mentioned. If one goes to see a medical practitioner for a mental health question, one is well aware of one's emotional state: it is the reason why they visit a professional. Mental health check-ups are rarely realized: we do not seem to go to the psychiatrist in the same way we go to the dentist, for an annual exam. The patient's self-awareness is at the core of psychiatry. Furthermore, it is also the psychiatrist role to discern whether the individual is experiencing this emotional fixity, despite them not being fully aware of it.

5 CONCLUSION

In this thesis, I tried to indicate the limitations of employing evolutionary principles to demarcate the normal from the pathological in psychiatry, especially in the case of depression. We cannot ground normality in evolved functions for two main reasons: first, because a states' expression in conformity to its evolved function cannot guarantee its normality; second, because there is a discrepancy between the way our emotions operate now, and how they did at the time they were selected for. More broadly, a functionalist approach – whether it appeals to evolved functions or not – seems to be unfit for the definition of normality in psychiatry, because of our ability to constantly transgress and modify the typical way our emotions are expressed. I then tried to argue that it might be beneficial to abandon a framework based on the normality-abnormality distinction in psychiatric diagnosis, for an intervention-based framework to avoid those issues. However, this conclusion does not entail that evolutionary theory has nothing to offer to the understanding of depression: as Nesse has claimed, if pathological depression does originate from low mood, understanding that there are sub-types of low mood, could potentially help in discerning different corresponding sub-types of depression: evolutionary theory can provide explanatory hypotheses – which would need to be further developed by psychological, biological and neurological studies – suggesting possible ways of classifying subtypes of depression. To view evolutionary theory as a means that could help research development on the classification problem of depression does not need to entail using evolutionary theory to demarcate normal from pathological instances of low mood. In other words, we can use low mood's evolutionary history to shed light on some aspects of current forms of depression without assuming that its evolutionary history provides us with the determining explanatory factor of how it operates today.

REFERENCES

- American Psychiatric Association. (1980). *Diagnostic and Statistical Manual of Mental Disorders, Third Edition*. Washington, D.C.: American Psychiatric Press.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, 5th Edition*. Washington: American Psychiatric Publishing.
- Bigelow, J., & Pargetter, R. (1987). Functions. *Journal of Philosophy*, 84(4), 181-196.
- Buller, D. J. (2005). Evolutionary psychology: the emperor's new paradigm. *Trends in Cognitive Sciences*, 9(6), 277-283.
- Burkhouse, K., Jacobs, R., Peters, A., & et al. (2017). Neural correlates of rumination in adolescents with remitted Major Depressive Disorder and healthy controls. *Cognitive Affective Behavioral Neuroscience*, 17, 394-405.
- Buss, D. M. (2013). Sexual jealousy. *Psychological Topics*, 22, 155-182.
- Canguilhem, G. (1991). *The Normal and The Pathological*. (C. R. Fawcett, & R. S. Cohen, Trans.) New York: Zone Books.
- Castiglioni, M., & Laudisa, F. (2015). Toward psychiatry as a 'human' science of mind. The case of depressive disorders in DSM-5. *Frontiers in Psychology*, 5, 1-12.
- Cudzik, M., Soroka, E., & Olajossy, M. (2019). Dissociative identity disorder as a wide range of defense mechanisms in children with a history of early childhood trauma. *Current Problems of Psychiatry*, 20(2), 1-13.
- Cummins, R. (1975). Functional analysis. *Journal of Philosophy*, 72, 741-765.
- Doyle, C., Swain, W., Edwald, H. S., & Edwald, P. W. (2019). Inflammation, infection and depression: an Evolutionary perspective. *Evolutionary Human Sciences*, 1(14), 1-19.
- First, M., Bhat, V., Adler, D., Dixon, L., Goldman, B., Koh, S., . . . Siris, S. (2014). How do clinicians actually use the Diagnostic and Statistical Manual of Mental Disorders in clinical practice and why we need to know more. *The Journal of Nervous and Mental Diseases*, 202, 841-844.
- Fried, E. I., & Nesse, R. M. (2015). Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. *BMC Medicine*, 13(72), 1-11.
- Goldberg, D. (2011). The heterogeneity of "major depression". *World Psychiatry*, 10, 226-228.

- Gould, S. J., & Lewontin, R. C. (1979). The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. *Proceedings of the Royal Society of London series B*, 205, 581-598.
- Hasin, D., Sarvet, A., Meyers, J., & et al. (2018). Epidemiology of adult DSM-5 Major Depressive Disorder and Its specifiers in the United States. *JAMA Psychiatry*, 75(4), 336-346.
- Hidaka, B. (2012). Depression as a disease of modernity: explanations for increasing prevalence. *Journal of Affective Disorders*, 140, 205-214.
- Horwitz, A., & Wakefield, J. (2007). *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder*. New York: Oxford University Press.
- Keller, M. C., & Nesse, R. M. (2005). Is low mood an adaptation? Evidence for subtypes with symptoms that match precipitants. *Journal of Affective Disorders*, 86, 27-35.
- Kendler, K. S., Myers, J., & Zisook, S. (2008). Does bereavement-related major depression differ from major depression associated with other stressful life events? *American Journal of Psychiatry*, 165, 1449-14955.
- Klinger, E. (1975). Consequences of commitment to and disengagement from incentives. *Psychological Review*, 82, 1-25.
- Lloyd, E., Wilson, D. S., & Sober, E. (2011). Evolutionary mismatch and what to do about it: A basic tutorial. <https://evolution-institute.org/wp-content/uploads/2015/08/Mismatch-Sept-24-2011.pdf>.
- Luoto, S. (2020). Did prosociality drive the evolution of homosexuality? *Archives of Sexual Behavior*, 49, 2239-2244.
- Magree, V. (2002). Normal and abnormal: Georges Canguilhem and the question of mental pathology. *Philosophy, Psychiatry, & Psychology*, 9(4), 299-312.
- Maj, M. (2016). The Continuum of depressive states in the population and the differential diagnosis between “normal” sadness. In J. Wakefield, & S. Demazeux (Eds.), *Sadness or Depression? International Perspectives on the Depression Epidemic and its Meaning* (Vol. 15, pp. 29-38). Dordrecht, Heidelberg, New York, London: Springer, History, Philosophy and Theory of the Life Sciences.
- Nesse, R. M. (2001). On the difficulty of defining disease: A Darwinian perspective. *Medicine, Health Care and Philosophy*, 4(1), 37-46.

- Nesse, R. M. (2009). Explaining Depression: Neuroscience is not enough, evolution is essential. In C. Pariante, R. Nesse, R. Nutt, & D. Wolpert (Eds.), *Understanding Depression: A Translational Approach* (pp. 17-35). Oxford: Oxford University Press.
- Nesse, R. M. (2009b). Evolution at 150: Time for truly biological psychiatry. *The British Journal of Psychiatry*, 195(6), 471-472.
- Nesse, R. M. (2015). Evolutionary Psychology and Mental Health. In *The Handbook of Evolutionary Psychology* (pp. 903-927). Hoboken, USA: John Wiley & Sons, Inc.
- Nesse, R. M. (2017). Evolutionary foundations for psychiatric research and practice. In V. A. B. J. Sadock, *Kaplan's and Sadock's Comprehensive Textbook of Psychiatry, 10th Edition, Volume One* (pp. 769-780). Philadelphia: Wolters Kluwer.
- Nesse, R. M., & Jackson, E. D. (2011). Evolutionary foundations for psychiatric diagnosis: Making DSM-V valid. In A. P. De Block A., *Maladapting Minds: Philosophy, Psychiatry, and Evolutionary Theory* (pp. 167-191). Oxford: Oxford University Press.
- Nesse, R. M., & Stein, D. (2012). Towards a genuinely medical model For psychiatric nosology. *BCM Medicine*, 10(5), 1-9.
- Ormel, J., & et al. (2020). What is needed to eradicate the depression epidemic, and why. *Mental Health and Prevention*, 17, 1-11.
- Pradeu, T. (2011). Philosophie de la biologie. In A. Barberousse, D. Bonnay, & M. Cozic (Eds.), *Précis de Philosophie des Sciences* (pp. 378-403). Paris: Vuibert.
- Sterelny, K., & Griffiths, P. (1999). *Sex and Death, An Introduction to Philosophy of Biology*. Chicago and London: The University of Chicago Press.
- Summerfield, D. (2006). Depression: epidemic or pseudo-epidemic. *Journal of the Royal Society of Medicine*, 164(3), 161-162.
- Tully, I. (2019). Demarcating depression. *Ratio*, 32(2), 114-121.
- Valles, S. A. (2010). The mystery of the mystery of common genetic diseases. *Biology & Philosophy*, 25, 183-201.
- Wakefield, J., & Demazeux, S. (2016). Introduction: Depression, one and many. In *Sadness or Depression* (Vol. 15, pp. 1-16). Dordrecht, Heidelberg, New York, London: Springer, International Perspectives on the Depression Epidemic and its Meaning, History, Philosophy and Theory of the Life Sciences.

- Wrosch, C., Miller, M., & Craver, C. (2003). Adaptive self-regulation of unattainable goals: goal disengagement, goal re-engagement, and subjective well-being. *Personality and Social Psychology Bulletin of the Menninger Clinic*, 29, 1494-1508.
- Zisook, S., & Shuchter, S. R. (1991). Depression through the first year after the death of a spouse. *The American Journal of Psychiatry*, 148(10), 1346-1352.