

# A new interpretation of the modern two-pronged tests for insanity

## Why legal insanity should not be a ‘status defense’

*Johannes Bijlsma*

### 1 Introduction

The case against Anders Breivik has given fresh impetus to the debate about the nature and proper scope of the insanity defense. In large part, this is due to the specific approach of the Norwegian law on insanity. The mere fact that someone is ‘psychotic or unconscious’ merits an acquittal by reason of insanity under paragraph 44 of the Norwegian penal code. This is in contrast with most other modern tests for insanity. In addition to the prerequisite that the defendant suffers from a mental disease, these tests stipulate (in various wordings) a second, two-pronged condition: the illness must have affected the defendant’s appreciation of the wrongfulness of the act or his ability to act in conformity with the law.

Michael Moore has argued extensively that this prevailing modern approach of insanity is wrongheaded and that the basic tenets of the Norwegian test as well as those of some older tests are correct by acknowledging that a finding of mental illness alone merits an acquittal by reason of insanity. Mental illness – understood by Moore in the limited commonsense notion of being ‘crazy’ – should be excusing in itself. No further conditions should be required.

If Moore is right, then current laws on insanity are wrong in most jurisdictions. Therefore, Moore’s challenging critique of the prevailing modern tests of insanity and his alternative ‘status approach’ merit a critical re-examination of these modern tests. In this paper I will present a new rationalization of the modern tests for insanity that evades Moore’s criticism but that also captures better why mentally disordered defendants should be acquitted than the status approach.

I will start with an outline of Moore’s criticism of the modern tests and his status approach (section 2). Then I will argue that the status approach of insanity is either under- or overinclusive (section 3). In the main part of the paper I will develop a new interpretation of the modern tests of insanity that hinges on the question whether or not a legally relevant difference between the ‘normal’ defendant and the mentally ill defendant can be established.<sup>1</sup> In sections 9 and 10 I will argue that the approach of insanity espoused in this paper can be accommodated to different legal systems and that the modern tests are not – contrary to

1 ‘Normal defendant’ is used in this paper to indicate defendants who do not suffer from mental disorder, i.e. the defendant who meets criminal law’s presuppositions about normal psychological functioning.

Johannes Bijlsma

Moore's critique – underinclusive. I will conclude that the rationalization of the modern tests presented here is not only better suited to distinguish between responsible and non-responsible defendants, but also facilitates meaningful debate in the courtroom about the criminal responsibility of the defendant.

## 2 Moore on 'madness' and criminal responsibility

According to Moore, mentally disordered defendants are exempt from punishment because they are irrational. There is a close link between people's everyday understanding of their own and other people's behavior ('folk psychology') and the concept of rationality Moore refers to. People understand human behavior by referring to the mental states – beliefs and desires – that rationalize an action:

'Here are two rather homely examples: suppose we ask Smith why he is carrying his umbrella to lunch. He replies that he believes it is going to rain. Suppose we ask Jones why he went across the street a moment ago, and he replies it was because he wanted some tobacco from the store across the street. Each, by his answer, has rendered his behaviour intelligible by giving us the practical reasoning that led him to do what he did.'<sup>2</sup>

The explanation of behavior in terms of mental states means that a practical syllogism is applied, as was already acknowledged by Aristotle. The premises of the syllogism are the desires ('I do not want to get wet') and the beliefs ('It is going to rain') of the agent. The conclusion is an intention to act and the action based on the intention ('carry an umbrella to lunch'). Citing the beliefs and desires makes the action a rational thing to do, if the action logically follows from its premises. Conversely, an action is irrational if the action is not logically related to the premises. Such is the case if the actor mentioned in the example would have the desire to stay dry, the belief that taking the umbrella would fulfill this desire, and yet leave the umbrella in the umbrella stand. People assume that rational actors will act to further their desires in light of their beliefs.

There are several implicit presuppositions in the premises of the folk psychology as described by Moore. By making these presuppositions explicit Moore maps the various ways in which actions can be irrational.<sup>3</sup> Firstly, in cases of extreme mental illness there are no beliefs and desires that could possibly rationalize an action. The bodily movements during an epileptic fit, for example, do not appear to be related to the agent's beliefs or desires in any way. A second way for an action to be irrational, is when the agent decides to behave in a certain way and nothing stands in the way of this action, but still cannot make himself do what he has decided to do ('weakness of will'). Thirdly, mental states as such can be irrational. Sometimes it is irrational to have a certain desire: 'Wishing to expose one's

2 Michael S. Moore, *Placing Blame: A Theory of the Criminal Law* (Oxford: Oxford University Press, 2010), 603.

3 Moore, *Placing Blame*, 604-8.

green books to the sun, as it stands, is unintelligible for a person to want intrinsically.<sup>4</sup> A belief is irrational if there is a disproportion between the evidence supporting the belief and the degree of confidence with which the belief is held, as is the case when one believes he is made out of glass. Intentions are irrational either if the actor intends something he deems impossible to achieve or if the agent's intention is not embedded in a larger plan. The irrationality of mental states 'infects' the irrationality of the action based upon it. The outcome of the practical syllogism is no longer rational if its premises are irrational. Lastly, an action is irrational according to Moore if one is acting on inconsistent desires or on conflicting intentions.

So far, Moore has only spoken of the irrationality of *actions*. However, the irrationality of a certain action is not the reason mentally ill defendants should not be held criminally responsible for their actions. The mentally ill should not be punished, according to Moore, because they are *themselves* irrational. Moore thus equates mental disorder with irrationality: "To say of someone that he is mentally ill is to say that he is irrational."<sup>5</sup> The rationality of agents is a function of the rationality of their actions over time. The more often one acts irrational, the more severe one is mentally disordered. To be a responsible agent is to have the ability to behave in a rational way regularly over time. Only then is it possible to understand an agent as people understand themselves and others in their everyday lives and will they regard him as a moral agent.<sup>6</sup>

This leads Moore to the conclusion that mental illness itself is excusing: a defendant should not be punished if he was so irrational that he cannot be justly held responsible.<sup>7</sup> If someone is sufficiently irrational, he is mentally ill, and therefore not a moral agent. As with animals and young children, he cannot engage in practical reasoning the same way as average adult human beings can. (Moore repeatedly cites Manfred Bleuler's famous phrase that the truly 'crazy' are 'stranger to us than the birds in our gardens.'<sup>8</sup>) Therefore he cannot be the object of blame. The logical consequence of this position is that it is not necessary to inquire whether there is some kind of relationship – be it a causal one or a lack of appreciation of the wrongfulness of the conduct or a lack of control – between the mental disorder and the criminal act. The 'status' of being mentally ill is a sufficient condition to exempt the defendant from punishment. This makes the insanity defense, according to Moore, a true 'status defense.'<sup>9</sup>

At this point, one might wonder whether Moore would want to excuse every defendant diagnosed by a psychiatrist to be mentally ill. That surely would lead to a very broad scope of the insanity defense. Psychiatric classifications like the

4 Moore, *Placing Blame*, 606.

5 Moore, *Placing Blame*, 605.

6 Moore, *Placing Blame*, 608.

7 Moore, *Placing Blame*, 609.

8 Moore, *Placing Blame*, 605; Michael S. Moore, 'The Quest for a Responsible Responsibility Test: Norwegian Insanity Law after Breivik,' *Criminal Law and Philosophy* (2015): 678, 682, 686.

9 Moore, 'The Quest for a Responsible Responsibility Test,' 680-83.

Johannes Bijlsma

*Diagnostic and Statistical Manual of Mental Disorders* classify a wide range of conditions as mental disorder. Therefore, a large proportion of criminal acts would be excused if mental illness were to be equated with the conditions that are recognized as mental disorder in psychiatry. Indeed, Moore's definition of mental illness for legal purposes is much narrower. When Moore speaks of mental illness, he is referring to a commonsense notion of mental illness that is expressed in unscientific terms like 'madness' or 'craziness.' Only very serious conditions – for example hallucinations and delusions – qualify as mental illness in this commonsense notion.<sup>10</sup>

### 3 'Madness' as a 'status defense' is either under- or overinclusive

As was explained above, Moore defines the rationality of a particular agent as a function of the rationality of his actions over time. The more often one acts irrational, the more severe the mental illness. It is safe to assume that under Moore's definition one single irrational act is not enough to acquire the status of being 'mad.' Furthermore, a person whose actions are *always* irrational can be presumed to be 'crazy' under Moore's definition. Somewhere in between must be the tipping point where the agent becomes so severely irrational that he acquires the status of being 'mad' and loses the status of responsible agent.

Moore's definition of 'madness' as a function of the rationality of an actor's actions over time, raises questions about the amount of time needed to constitute 'madness.' Some mental disorders are very severe, but rather short-lived. A delusion disrupts the proportionality between the evidence supporting certain beliefs of a defendant and the degree of confidence with which the beliefs are held, in such a way that all actions based on this belief become irrational by Moore's definition. By any definition of mental disorder, delusions would qualify as serious illness.<sup>11</sup> Suppose that a persecutorial delusion lasts for a week (meeting the DSM 5-threshold for brief psychotic disorder<sup>12</sup>) and that most actions of the defendant during that week are based on the false belief of being persecuted. During that week the defendant kills somebody he believes is plotting to murder him. The delusion does not recur. Has this defendant acquired the status of being 'mad'? Is a week (or a day, or an hour) of suffering from delusions enough to be 'mad' under Moore's definition? If not, the defendant cannot successfully plead insanity. However, his *actions* based on the delusion are indisputably irrational under Moore's definition. Should this defendant be denied exemption from punishment for actions that are clearly irrational due to a serious mental disorder and irrationality – as Moore proposes being the test for insanity – merely because the defendant has – presumably – not been haunted long enough by delusions to acquire the status of being mad? It seems rather arbitrary to deny a defendant

10 Moore, 'The Quest for a Responsible Responsibility Test,' 683-87.

11 See also Moore, 'The Quest for a Responsible Responsibility Test,' 657, 685-86.

12 *Diagnostic and Statistical Manual of Mental Disorders*. Fifth edition (Arlington: American Psychiatric Association, 2013), 94.

recourse to the insanity defense *only* because he has not been ill for a long enough period.

Conversely, if the defendant should be considered 'mad,' when does he regain the status of morally responsible agent? Because the irrationality of a particular agent is a function of the irrationality of his actions over time, 'madness' has a certain durability. This durability causes the 'madness' to persist for at least a while after the symptoms of irrationality have disappeared. However, it seems problematic to acquit a defendant by reason of insanity if the symptoms of his illness are not manifest.

So far, my interpretation of the status approach may have put too much emphasis on the durability of 'madness.' It may very well be that the status of being 'crazy' is not as rigid as suggested above. Even then, however, the same problem seems to exist. The irrationality of mental states 'infects' only the actions based upon those mental states, and that allows for actions to be not irrational according to Moore's definition even when the defendant is suffering from delusions. After all, it is not likely that a delusion 'infects' *every* action the defendant performs. The defendant may for example steal something out of greed, not because he deludedly thinks somebody is chasing him. Since the defendant is 'mad' while he steals, he should be acquitted by reason of insanity under Moore's status test.<sup>13</sup>

Moore's leap from the irrationality of single actions to the irrationality of agents thus has as a consequence that irrational agents are not responsible for actions that are rational according to Moore's concept of practical reason. It seems difficult to justify an acquittal of a defendant for an act that is *not* irrational if – as proposed by Moore – irrationality is the test for insanity. Moore does not explain why the irrationality of other actions of the defendant should have a bearing on his legal responsibility for a criminal act that is not irrational (or vice versa), except for the circular argument that the insanity defense is a status defense. Therefore, Moore's test seems to be overinclusive in this respect or he has to accept that if a defendant is still capable of acting rational sometimes, he cannot be 'mad.'

However, if *every* action the defendant performs has to be irrational, the test for insanity would become *very* strict. Most cases in which the defendant suffers from delusions or hallucinations would not lead to insanity since even those defendants will act rational every now and then. Irrational actions performed on the deluded belief that one is being persecuted, would not constitute a defense if the defendant has been acting rational in some other instances. That would seem to make the test underinclusive. Why would the rationality of other actions preclude a defense for an action that clearly is irrational due to a serious mental disease?

13 Moore, 'The Quest for a Responsible Responsibility Test,' 681.

Johannes Bijlsma

Under this strict interpretation, only very serious disorders of the rational capacities, like profound developmental disorder or advanced cases of dementia,<sup>14</sup> would qualify as ‘madness,’ since it is conceivable that these disorders render every action irrational. Moreover, most (or even all) of these cases would not be treated under the insanity defense, because the defendant did not act with the required *mens rea*. Under the strict interpretation of the status test a test for insanity may well prove to be superfluous.

#### 4 A new rationale for the modern two-pronged approaches to the insanity defense

Moore is not convinced by the turn the law of insanity has taken in most Western legal systems for (roughly) the last century and a half. Modern tests for insanity supplement the condition that the defendant has to suffer from mental disorder with a second condition for insanity, mostly consisting out of two prongs. The first prong is usually (in various wordings) that the defendant failed to appreciate the illegality or wrongfulness of his conduct in order for him to be excused. To this cognitive prong a volitive prong is usually added, excusing the defendant (in various wordings) if he could not control his conduct.<sup>15</sup> This last prong is controversial in Anglo-American jurisdictions.<sup>16</sup> In these tests the focus is on the act, not on the moral status of the defendant.

Moore thinks this modern turn to a two-pronged approach is wrongheaded, because there is no need for the separate condition that the defendant suffers from mental disorder if ignorance or mistake of law or loss of control excuses defendants. ‘What excusing work is left to be done by mental illness?’ Moore asks rhetorically.<sup>17</sup> This line of reasoning is taken to its ultimate conclusion by Slobogin, who argues that the insanity defense can be abolished as a separate excuse because it can be integrated in other doctrines like *mens rea* and putative justification. Mental disorder should, according to Slobogin, excuse only if, for example, the *mens rea* element of the crime cannot be proven due to the disorder or when a disorder leads to the mistaken belief that the conduct was justified. The regular doctrines for assessing criminal responsibility cover all situations in which a disorder should excuse. The disorder does no independent ‘excusing work’ under Slobogin’s integrationist test of insanity.<sup>18</sup>

14 Cf. *DSM-5*, 36.

15 E.g., par. 122-1 Code Pénal (France); par. 4.01-1 Model Penal Code (U.S.A.); par. 31-1 sub a Rome Statute (International Criminal Court); par. 20 Strafgesetzbuch (Germany).

16 See below, section 7.

17 Moore, ‘The Quest for a Responsible Responsibility Test,’ 662.

18 Christopher Slobogin, ‘An End to Insanity: Recasting the Role of Mental Disability in Criminal Cases,’ *Virginia Law Review* (2000): 1199-1247; Christopher Slobogin, ‘A Defense of the Integrationist Test as a Replacement for the Special Defense of Insanity,’ *Texas Tech Law Review* (2009): 523-42.

Conversely, Moore argues that if mental illness is doing part of the ‘excusing work,’ the question arises why the mentally disordered defendant is excused under circumstances that otherwise would not excuse. Ignorance or mistake of law in general do not excuse, after all. The same goes for loss of control. One who is easily provoked is not excused from violent behavior, but when loss of control is a result of mental illness the second prong does excuse the defendant. Thus, it seems mental illness is doing at least some of the excusing work, since excusing conditions that do not normally lead to an acquittal do so if the defendant is suffering from mental disorder.

Indeed, as was described above, *all* excusing work should be done by mental illness, according to Moore, since ‘madness’ bereaves the defendant from moral agency. Moore’s approach is the exact mirror image of Slobogin’s. Slobogin argues that from a legal perspective there is nothing very special about mental illness and that ‘normal’ legal doctrines can be used to establish the criminal responsibility of the disordered defendant, while Moore’s key point is that mental illness is a defense *sui generis*, not to be compared with the other excuses. Older tests for insanity and the current Norwegian test for insanity do acknowledge this and should therefore be preferred over the currently prevailing two-pronged approaches to insanity.<sup>19</sup>

If Moore is right, the law on insanity in most Western legal systems has taken a wrong turn and should be replaced by a status approach. As was argued above, his status approach to insanity is quite crude: by its own logic it seems to be either under- or overinclusive. Moore’s arguments against the two-pronged approach boil down to the fact that the relationship between the requirement of mental disorder and both prongs is unclear. However, there is a way to think of the relationship between mental disorder and both prongs of the prevailing modern tests that on the one hand counters Moore’s objections and on the other hand avoids the pitfalls of the status test.

Since these modern tests exist in many jurisdictions, I will offer no analysis of how the interpretation expounded here relates to positive law on insanity in different legal systems. My main goal in the next sections is to offer a general conceptualization of the two-pronged approach that avoids Moore’s criticism. In a separate section I will argue that this interpretation can be accommodated to different legal systems.<sup>20</sup>

## 5 On the relation between mental disorder and both prongs

Moore’s link between the psychological presuppositions underpinning the criminal law and people’s everyday understanding of human psychological functioning

19 Moore, ‘The Quest for a Responsible Responsibility Test,’ 678-80.

20 See below, section 9.

is widely acknowledged in scholarly literature.<sup>21</sup> These presuppositions are often termed ‘folk psychology.’ It is criminal law’s basic assumption that psychologically sound citizens can be held accountable for their actions because of their ability to engage in practical reasoning. In exceptional cases this condition for criminal responsibility is not met: in one or more important respects the defendant did not meet the folk psychological criteria for sound psychological functioning while breaking the law and has to be considered mentally ill. Obviously, if basic assumptions for criminal responsibility are not met, the criminal responsibility of the defendant for the act is in dispute. Whether or not a defendant is to be acquitted on this ground is the domain of the insanity defense.

The first requirement for insanity under the two-pronged approach is the establishment of mental disorder at the time of the act. A straightforward legal definition of mental disorder would be that a defendant suffers from mental disorder if the criminal law’s presuppositions about normal psychological functioning have not been met.

This is the approach Moore takes as the starting point of his reasoning where he discusses the (ir)rationality of actions as a result of an impaired ability to engage in practical reasoning. Moore, however, is not prepared to accept merely an impaired ability to engage in practical reasoning at the time of the act as an instance of mental disorder in a legal sense. The actor must have acted irrational sufficiently often over time to be considered ‘crazy.’

Moore is more or less forced to take this turn, because he considers mental illness itself excusing. If he were not to limit his concept of mental illness – ‘craziness’ – to the most severe cases, the scope of the insanity defense would become very broad. *Every* irrational action – every crime based on irrational beliefs and desires (etcetera) – would then have to be excused.

Under the two-pronged approach there is no *a priori* need to limit the concept of mental illness to the most severe cases, since the scope of the defense is not only determined by the concept of mental illness, but narrowed by the second requirement: one of both prongs for insanity needs to be fulfilled. Therefore, under the two-pronged approach, it can be agreed with Moore that to be mentally ill in terms of the insanity defense is to act sufficiently irrational at the time of the act as not to meet the folk psychological presuppositions underpinning criminal law, without further limiting this concept to conditions that are durable over time and to most severe cases.

Once it has been established that the defendant does not meet the folk psychological criteria for sound psychological functioning, it must be assessed whether

21 See, e.g., Andrew E. Lelling, ‘Eliminative Materialism, Neuroscience and the Criminal Law,’ *University of Pennsylvania Law Review* (1993): 1477-83; Stephen J. Morse, ‘Determinism and the Death of Folk Psychology: Two Challenges to Responsibility from Neuroscience,’ *Minnesota Journal of Law, Science and Technology* (2008): 4-6; Katrina L. Sifferd, ‘In Defense of the Use of Commonsense Psychology in the Criminal Law,’ *Law and Philosophy* (2006): 571-612.

one of the two prongs for insanity has been fulfilled. With a view to Moore's criticism, why should this question be raised at all? Indeed, ignorance or mistake of law and loss of control are in general no excusing conditions to 'normal' defendants. Why should mentally disordered defendants be excused under circumstances that do not generally excuse?

Fulfillment of the first condition of mental disorder entails the determination that the defendant did not meet the presuppositions of sound psychological functioning that are presupposed for criminal responsibility. However, the mere fact that this is the case does not merit an acquittal by reason of insanity under the two-pronged approach. The determination that the defendant suffers from mental disorder does not necessarily imply that the disorder had a legally relevant bearing on the conduct of the defendant. The first condition of mental disorder therefore can be regarded as meriting a further examination as to whether the defendant can be held criminally responsible. The legal relevance of mental disorder as an excusing condition can only be that the disorder constitutes a legally relevant difference with the 'normal' defendant, meriting an acquittal in circumstances that otherwise would not constitute an excuse. This can be grounded in the uncontroversial general principle that to merit a different treatment, a relevant difference has to exist.

A legally relevant difference is present if the mentally disturbed defendant cannot be held to the same normative standards as 'normal' defendants due to mental disorder. Both prongs – the second condition for insanity – can be viewed as mapping those cases in which a legally relevant difference with the 'normal' defendant is present and thus when an acquittal is warranted. In order to understand why mentally disordered defendants should be acquitted it should be made intelligible why exactly the defendant should be excused under the circumstances described by both prongs.

## 6 Legally relevant differences and the cognitive prong

Mistake of law in general does not excuse. However, as Moore acknowledges, under exceptional conditions to have no knowledge of the wrongfulness of conduct is an excuse available to all defendants. The first prong of the modern insanity tests therefore does not necessarily introduce a defense at odds with the maxim *ignorantia juris neminem excusat*, since this rule already allows for exceptions. The question is rather why a special provision for mentally disturbed defendants is needed in addition to the excuses available to every defendant.

One way for a 'normal' defendant to be excused for having no knowledge of the wrongfulness of his conduct is to act in putative justification.<sup>22</sup> Putative justification grants an excuse to a defendant who, acting under the mistaken belief that the facts were such that his conduct was justified, fulfills the elements of an

22 See also Moore, 'The Quest for a Responsible Responsibility Test,' 658-59, 662.

offence. Usually it is required that the mistake of fact was reasonable.<sup>23</sup> Two conditions for an acquittal on the ground of putative justification therefore need to be established: (1) if the mistaken belief would have been true, the conditions of a justification like self-defense or duress would have been fulfilled and (2) this factual error was reasonable.

The condition of reasonableness of the factual error entails that it is not enough for the defendant to merely have the belief that his conduct was justified, but also that a 'reasonable citizen' in the same circumstances would have made the mistake as well. The actor is not allowed to just act on his belief, he must – within reasonable boundaries – assess whether or not his belief is true. If a reasonable person would have concluded that the belief held by the defendant was erroneous, the defendant is not excused, even though he believed his conduct was justified.

The requirement of the reasonableness of the factual error implies that the defendant is able to critically reflect upon his own factual beliefs and to correct them when confronted with conflicting evidence. As Moore has shown, the law indeed presupposes the average citizen to possess this psychological capacity. Therefore, a requirement of reasonableness of a mistake of fact is justified in general.

However, mental illness can affect this psychological capacity. A delusion, for example, can be defined as an unwarranted belief that cannot be corrected in the light of conflicting evidence.<sup>24</sup> A deluded defendant acting in presumed self-defense on the mistaken belief that he is (about to be) attacked, will not change this belief even if he were to be confronted with overwhelming evidence to the contrary. By definition, a reasonable person would know the deluded belief to be erroneous. Because of the unreasonableness of his belief, the mentally disordered defendant cannot rely on a defense of putative justification since he does not meet the second condition. A conviction, however, would clearly not be justified, since the psychological capacity presupposed by the second condition of putative justification – the capacity to reflect on beliefs and to correct them if necessary – does not hold with regard to this defendant with respect to the act committed by him.<sup>25</sup> It is therefore necessary to have an excuse available to defendants who cannot be held to the same normative standards – in this case: the demand to

23 Jeroen Blomsma, *Mens Rea and Defences in European Criminal Law* (Cambridge: Intersentia, 2012), 299-309; Wayne R. LaFare, *Substantive Criminal Law* (St. Paul: West, 2017), par. 10.1 (d), 10.4 (c). Some jurisdictions do not have this requirement of reasonableness for certain putative defenses. See below, section 9. Putative justification should be distinguished from mistake as to whether the elements of the offense were fulfilled, which might negate *mens rea*. When acting under putative justification, the defendant does not err about fulfilling the elements of the offense, but mistakenly thinks the offense was justified.

24 *DSM-5*, 87.

25 See for this line of reasoning *People v. Leeds*, 240 Cal.App. 4th 822 (2015), 829-33. Cf. Victor Tadros, *Criminal Responsibility* (Oxford: Oxford University Press, 2005), 336.

critically reflect on beliefs – as the ‘normal’ citizen. The first prong of the modern tests for insanity can be interpreted as to provide for just that.

The first prong allows the defense only to defendants who did not know they were acting contrary to the law. This limitation makes sense. Apart from mistakes of fact that negate the *mens rea* element of the offense,<sup>26</sup> mistakes of fact that do not affect the knowledge of the defendant of the wrongfulness of the act are not legally relevant. The delusion is primarily relevant vis-à-vis the condition of the reasonableness of the belief, not with respect to the first condition of putative justification. In other words: the second condition of putative justification cannot be applied in case of delusions, but the first condition can in principle be maintained, since there is no legally relevant difference that would allow for not applying that condition. However, standards governing the second limb of putative justification like subsidiarity and proportionality may be *lowered* if that would be necessary in view of the particular capacities of the defendant at the time of the act.

Not only do mistakes of fact leading to putative justification affect the knowledge of any defendant of the wrongfulness of the act, mistake of law can do so as well. In case of mistake of law, the defendant is aware of the factual circumstances of his conduct, but he errs (reasonably or even unavoidable) regarding the existence or the scope of the legal norm that applies to his conduct. Section 2.04 (3) Model Penal Code, for instance, allows for an excuse of mistake of law if the law was not published or reasonably made available, or if the defendant relied on an official (but erroneous) statement of the law.

The rationale of this excuse is not so much that every citizen (supposedly) knows what the law is, but rather that every citizen is able to inform himself on what the law is.<sup>27</sup> Clearly, a duty to inform oneself on the law cannot be imposed if the law is not published or reasonably made available. Likewise, one cannot be held accountable if one exercised the duty to inform oneself on the state of the law, but an official gives an incorrect statement on the state of the law.

Underlying this rationale is the human capacity of practical reason, as described by Moore. ‘Normal’ citizens of a certain age are able to understand what legal norms are, what their substance is and are able to apply those norms on their (future) conduct. That way, they are able to understand whether their conduct was allowed or prohibited, and adjust their behavior accordingly.

With some mental disorders, the intellectual capacity of the defendant is affected in such a way that it no longer is reasonable to hold the defendant to the same

26 See footnote 25.

27 See also Blomsma, *Mens rea and defences in European Criminal Law*, 469-74; LaFave, *Substantive Criminal Law*, par. 5.6(d). Husak has recently argued that blameworthiness of actors for ignorance of the law depends on how culpable they are with regard to their ignorance. Only the actor who is fully aware of the prohibited nature of his conduct is fully responsible. (Douglas Husak, *Ignorance of Law: A Philosophical Inquiry* [Oxford: Oxford University Press, 2016], 136-90.)

Johannes Bijlsma

norms as ‘normal’ citizens. A defendant suffering from a severe intellectual development disorder or dementia may not be able to understand the concept of legal norms, to understand their substance or to apply those norms to future conduct.<sup>28</sup> In such a case, it cannot be considered reasonable to ask of the defendant to inform himself on the state of the law and to apply the law, as would be demanded from other citizens, for even if he were to try to inform himself, he would not be able to understand the illegality of his contemplated conduct. Therefore, in this case, there is a legally relevant difference with the ‘normal’ defendant, allowing for an excuse under circumstances that would not normally be excusable.

Since ignorance of the law in general does not excuse, it is not enough for the mentally disordered defendant to in fact not understand the illegality of his conduct. For a successful defense of insanity he should not know in fact that his conduct constituted a crime *and* he should not be able to understand the wrongfulness of his act if he were to inform himself about the scope of the law. Only in that case, an exception for mentally disturbed defendants to the general rule of *ignorantia juris neminem excusat* is warranted.

Note that it is not the mental disorder as such that does the ‘excusing work,’ since not only the mental capacities of the defendant in general have to be assessed. It has to be examined whether or not the disorder made it impossible for the defendant to appreciate the wrongfulness of his behavior. This should be judged taking the subjective rational capacities of the defendant in account, not according to objective reasonableness standards. A disorder may, for example, not stand in the way of understanding the illegality of conceptually easy to grasp crimes like murder, but do hamper the appreciation of conceptually more complicated crimes.

## 7 Legally relevant differences and the control prong

So far, only instances were considered in which an acquittal by reason of insanity should be permitted because the first prong has been met. The second prong excuses the mentally disturbed defendant if he could not control his conduct to conform to the law.

First of all, it should be noted that this prong is somewhat controversial, especially in Anglo-American jurisdictions. Most notably England and Wales,<sup>29</sup> but also quite a few U.S.-jurisdictions, do not acknowledge lack of control as a ground for insanity. Continental jurisdictions, the Statute of Rome of the International Criminal Court and other U.S.-jurisdictions, however, do add a control test to a cognitive prong.

28 One of the characteristics of severe intellectual development disorder is: ‘The individual generally has little understanding of written language or of concepts involving numbers, quantity, time, and money’ (*DSM-5*, 36).

29 David Ormerod, *Smith and Hogan’s Criminal Law* (Oxford: Oxford University Press, 2008), 304.

Moore argues that, like ignorance of the law, loss of control in general does not excuse, and therefore, it is not clear why defendants should be excused if their loss of control was caused by mental illness if it were not mental illness itself that does the 'excusing work.' However, like the first prong of the two-pronged approaches to insanity, the second prong must imply that – in order to excuse the defendant – somehow there must be a legally relevant difference between a 'normal' defendant losing control and a loss of control caused by mental disorder that amounts to an acquittal by reason of insanity. If that cannot be established, there would be no reason to treat the mentally disordered defendant differently from any other defendant.

To the general rule that loss of control is no excuse for criminal conduct there are exceptions as well. Excuses available to every defendant show this. The excuse of self-defense-excess in Continental jurisdictions for example is available to defendants who exceed the boundaries of justificatory self-defense because of a forceful affect – like anger or indignation – caused by the attack.<sup>30</sup> Loss of control because of anger or indignation in general does not excuse, but in the particular circumstances of an unlawful attack it may.<sup>31</sup> A more general recognition of loss of control as an excuse is the defense of duress. To succeed, a defendant must have acted under external pressure – for example a threat – and could not reasonably be expected to resist that pressure.<sup>32</sup>

The second limb of this test makes clear that not every external pressure pursuant to which the defendant acted can be a ground for excuse, because then every instance of loss of control due to anger or impulsively acting on desires would exculpate. An important function of criminal law is precisely to deter citizens from acting on their impulses. They are expected to resist the desire to take something that does not belong to them, for example. It is possible to ask of citizens to conform their conduct to the law, because they are able to act to further their desires in the light of their beliefs.<sup>33</sup> Since criminal law's presupposition is that 'normal' citizens are able to understand the (il)legality of their intended conduct, they can be asked to shape their actions in the light of this belief.

Mental illness can affect the general psychological ability to shape actions in light of beliefs and desires. However, the mere establishment of mental disorder under the two-pronged approach is not enough to merit an acquittal. The disorder has to constitute a legally relevant difference with the 'normal' defendant that merits an acquittal in circumstances that otherwise would not excuse. To constitute a legally relevant difference with the 'normal' defendant, loss of control caused by

30 Blomsma, *Mens Rea and Defences in European Criminal Law*, 427-30.

31 In Anglo-American jurisdictions reasonable anger or indignation as a result of provocation may result in a partial defense for intentional homicide. (LaFave, *Substantive Criminal Law*, par. 15.2.)

32 Blomsma, *Mens Rea and Defences in European Criminal Law*, 442-44; LaFave, *Substantive Criminal Law*, par. 9.7.

33 Cf. Michael S. Moore, 'The Neuroscience of Volitional Excuse,' in Dennis Patterson & Michael S. Pardo (ed.), *Philosophical Foundations of Law and Neuroscience* (Oxford: Oxford University Press, 2016), 179-227.

Johannes Bijlsma

mental illness has to be qualitatively different from the kind of 'loss of control' a sound psychological functioning actor can experience when – for example – angry or impulsive.

This standard is not easily met. The loss of control experienced by normal defendants caused by forceful, but normal affects or impulsiveness can be quite real, but is generally not a ground for exculpation based on duress. Under the second prong, the disorder has to be of the severity of not being able to act in conformity with the belief that the conduct is illegal, an ability that 'normal' defendants – even when they are very angry or impulsive – are presumed to possess.

Some authors argue that pathological lack of control should not be an excusing condition at all, because of the empirical uncertainty regarding whether there actually exist mental disorders defined by a loss of control that can be distinguished clearly enough from 'normal,' non-excusing loss of control caused, for example, by anger.<sup>34</sup> Moore has recently mapped out an interesting method to operationalize folk psychological loss of control in empirical, neuroscientific terms, offering a plausible model to differentiate actors who *would* not control their behavior from actors who *could* not control behavior. His analysis, interestingly, seems to commit him at least partly to the model of insanity espoused here.

Moore carefully maps various ways in which there can be loss of control in folk psychological terms. In his view there is some room for a viable excuse of weakness of will. Two conditions need to be met: (1) the degree of self-control of an actor is lessened from what is normally demanded of persons generally and (2) it is not a moral defect in the person to have such lessened powers of self-control. Certain mental illnesses may meet both conditions and thus offer a volitional excuse, even if the mental illness does not amount to insanity.<sup>35</sup>

Moore derives both conditions for a volitional excuse from Tadros.<sup>36</sup> Tadros argues that standards of reasonableness in defenses like provocation should be adjusted to the personal characteristics of the defendant. One such characteristic may be mental disorder. In cases of domestic abuse, for example, it might not be reasonable to hold the defendant who killed her husband and is suffering from battered women syndrome to the same standards of self-control as 'normal' persons who are provoked. Other examples of possible volitional excuses Moore describes, may turn out to be forms of mental illness as well.<sup>37</sup> The normative

34 E.g., Stephen J. Morse, 'Culpability and Control,' *University of Pennsylvania Law Review* (1994): 1587-1660, 1619-34. See also Gerben Meynen, *Legal Insanity: Explorations in Psychiatry, Law, and Ethics* (Cham: Springer, 2016), 22-26.

35 Moore, 'The Neuroscience of Volitional Excuse,' 205.

36 Tadros, *Criminal Responsibility*, 349-58.

37 Kleptomania, for example, is characterized as follows: 'Individuals with kleptomania typically attempt to resist the impulse to steal, and they are aware that the act is wrong and senseless. The individual frequently fears being apprehended and often feels depressed or guilty about the thefts' (*DSM-5*, 478). This description may meet Moore's concept of loss of control as the inability to execute the right intention. (Cf. Moore, 'The Neuroscience of Volitional Excuse,' 201-4.)

standards, according to Tadros, should be tailored to the capacities of a particular defendant, not abandoned altogether.

This is similar to the approach of insanity taken here, with the important difference that the kind of cases are not considered under the law of provocation or other 'general' defenses, but under the insanity defense. Insanity under the two-pronged approach advocated here *by definition* implies that normative expectations of mentally disordered defendants should be tailored to their personal characteristics, sometimes not applying certain legal norms *at all*. In the next section I will argue why a special defense for the mentally ill is needed.

Not just any form of pathological loss of control will excuse, however. The loss of control produced by the mental disorder has to cause the defendant to not being able to act in conformity with the belief that the conduct is illegal (and no prior fault should apply). If the defendant – taking into account the particular disorder he suffered from and the bearing the disorder has on the possible application of general legal norms – was able to conform his behavior to the law, then he is a responsible agent (when no other excuses apply). In the case of battered women syndrome, for example, it should be assessed whether or not the objective second limb of duress can be applied as it would apply to 'normal' defendants and if not, to what extent the defendant could be expected to resist the urge to kill her husband.

## 8 No legally relevant difference: pedophilia

So far, cases have been discussed that might constitute a legally relevant difference and thus possibly lead to an acquittal by reason of insanity. However, applying the same method, it is also possible to identify cases that do not excuse.

Pedophilia is sometimes thought to be a disorder that might excuse under the control prong.<sup>38</sup> Again, it should be asked what the possible legally relevant difference could be with the 'normal' defendant. Nothing in the *DSM*-description of pedophilia suggests that the urges that characterize this disorder differ markedly from non-pathologic sexual desires, affecting the psychological capacity to shape actions in the light of beliefs and desires.<sup>39</sup> The difference here with 'normal' defendants seems to be that there is something wrong with the desires themselves.

Moore calls these desires irrational, because they are directed at an end that is morally wrong to pursue. This implies that the desire to engage in sexual conduct with minors is not considered to be a desire within the 'normal' range of desires and thus can be considered a disorder in the folk psychological sense as described by Moore. Moore concedes that his analysis commits him to the existence of pathological immoral desires, but he is not willing to call pedophiles 'crazy' and

38 Slobogin, 'An End to Insanity,' 1225.

39 Cf. *DSM-5*, 697-700.

Johannes Bijlsma

excuse them on the ground of insanity, because doing so would be to collapse the distinction between the mad and the bad.<sup>40</sup>

The interpretation of the two-pronged approach advocated here, offers another explanation as to why pathological immoral desires should not exculpate. A pathological desire that is at odds with the criminal law does not constitute a legally relevant difference with the 'normal' defendant. *Every* defendant who acts with intent has a desire to engage in proscribed conduct. Acting on a – albeit pathological – desire to engage in sexual conduct with minors therefore does not constitute a legally relevant difference. Pedophilia can therefore be no ground for an acquittal by reason of insanity for statutory rape. The two-pronged approach is correct in not adding a prong that would allow for pathological desires to exculpate.

## 9 Why is a separate defense necessary?

The advantage of the interpretation of insanity advocated here is that it is possible to maintain objective reasonableness standards for excuses like self-defense or duress, while at the same time allowing for exceptions for the mentally ill. Keeping objective reasonableness standards is important, amongst others, for the normative force of legal norms. Citizens can be sure which norms the law will hold them to. That is not to say that personal, non-pathological characteristics should play no role at all in assessing excuses like self-defense or duress. However, conflating other excuses with the insanity defense would mean that it is sometimes necessary – as was pointed out above – to lower the standards that apply to 'normal' defendants and sometimes to not apply those standards at all. That should only be allowed under a separate defense, reserved for special categories of defendants: the young, the involuntarily intoxicated and – indeed – the insanity defense.<sup>41</sup>

Standards for excuses vary across jurisdictions. It is important to note that the scope of the insanity defense under the approach advocated here, depends on the scope of the excuses that are available to 'normal' defendants in a certain jurisdiction. A jurisdiction that endorses a completely subjective approach to excuses like putative justification, ignorance of the law or duress does not have to provide a separate excuse to mentally disordered defendants. In such a jurisdiction, it would not be necessary to apply different norms to mentally disordered defendants than to 'normal' defendants, because the 'general' norms already allow for such tailoring.

40 Moore, *Placing Blame*, 606.

41 Cf. Moore, 'The Neuroscience of Volitional Excuse,' 205; Stephen J. Morse & Morris B. Hoffman, 'The Uneasy Entente between Legal Insanity and Mens Rea: Beyond Clark v. Arizona,' *The Journal of Criminal Law and Criminology* (2007): 1127.

I am not aware of legal systems where all excuses are governed by subjective standards only.<sup>42</sup> However, not all legal systems apply reasonableness standards all over the line. In England, for example, *honest* mistake on the facts is enough for an excuse of putative self-defense. Mistake regarding the *scope* of the law on self-defense should, however, be *reasonable*.<sup>43</sup> Since a delusion – by definition – is an honest mistake, it is in such jurisdictions not necessary to consider mistake on facts giving rise to self-defense due to insanity under the law of insanity. The first prong of the insanity defense, however, still is necessary to provide for a defense for *unreasonable* mistake of law due to mental disorder. In addition to that, unreasonable mistake on the facts leading to self-defense may in some jurisdictions give rise to negligence liability (e.g., in Germany).<sup>44</sup> In such a case, the first prong still is needed to account for legally relevant differences due to mental disorder.

The approach of insanity espoused here does not endorse a certain view on the proper scope of the other excuses. Rather, the proper scope of the insanity defense varies across legal systems, depending on the degree other excuses allow for a fair assessment of the capacities of the mentally disordered defendant. The more subjective those excuses are, the less ground is to be covered by the insanity defense. The interpretation defended here allows for accommodating for such differences and can thus be applied across jurisdictions.<sup>45</sup>

## 10 Are the modern tests underinclusive? Assessing M’Naghten

Moore’s main argument against the two-pronged approach to insanity is that there are defendants who fully understand the wrongfulness of their conduct and possess the ability to conform their behavior to the requirements of the law, but who clearly are crazy and therefore should be exempt from punishment. The limitations set forth by both prongs he deems to be too narrow.

As an example, Moore reflects on the famous case against M’Naghten. Daniel M’Naghten had the intent to kill the British Prime Minister Peel, but in fact shot Peel’s secretary. M’Naghten’s act was based on the delusion that Tories were persecuting him and out to hurt him. He knew what he was doing and knew that he acted contrary to the law; the facts as he believed them to be would not amount to a justification on the ground of self-defense in the absence of (putative) imminent danger. Neither was M’Naghten unable to conform his conduct to the requirements of the law. He surely was under the delusion that he had a hard choice to make, but the facts as he believed them to be would – if true – not

42 Slobogin fails to recognize this in his defense of his integrationist test. (Morse & Hoffman, ‘The Uneasy Entente between Legal Insanity and Mens Rea,’ 1127-29.)

43 Blomsma, *Mens Rea and Defences in European Criminal Law*, 304.

44 Blomsma, *Mens Rea and Defences in European Criminal Law*, 304.

45 In my dissertation I have used this conceptualization to develop an interpretation of a two-pronged test for insanity that is in line with principles of Dutch criminal law. (J. Bijlsma, *Stoornis en Strafwitsluiting* (Oisterwijk: Wolf Legal Publishers, 2016).) Dutch criminal law currently does not dictate a test for insanity.

Johannes Bijlsma

amount to a defense of duress, again because there was no imminent danger. The ironic consequence is that M’Naghten would flunk the test that bears his name, as well as the two-pronged tests for the insanity defense with a volitional prong.

Moore thinks this is not right. He appeals to a moral intuition that the deluded M’Naghten was ‘so crazy as not to be responsible’ and therefore should have been acquitted by reason of insanity.<sup>46</sup>

It is important to note that most severe cases of mental illness will not be considered under the insanity defense. In extreme cases, an acquittal will follow because either the defendant has not acted at all (e.g., epileptic seizures, dissociative states) or the *mens rea* element of the crime cannot be proven (e.g., certain delusions). Under the insanity defense, typically defendants are considered who are not wholly ‘stranger to us than the birds in our gardens.’ They have *acted* and they have acted with the required *mens rea*. Most of them are able to exercise practical reason. They may also possess the ability to apply general norms on their conduct and act accordingly.

In the cases of mental illness that are considered under the insanity defense the defendant thus possesses some of the psychological capacities that are presupposed for criminal responsibility, while lacking others.<sup>47</sup> M’Naghten is presumably lacking the capacity to correct his belief that he is being persecuted in the light of conflicting evidence. The lack of this capacity – his delusion – should lead to his acquittal according to Moore. Because of his delusion, under the status approach, M’Naghten is not responsible for *any* criminal act he commits while being ‘crazy.’ If, for example, it would have turned out that M’Naghten not only attempted to shoot Peel, but also robbed a bank because he was needy, he should be acquitted of the robbery under the status approach. Earlier in this paper I have argued that the status approach would be overinclusive if interpreted this way, because it allows for acquittals of actions that are not irrational (the supposed robbery).

Conversely, I am not convinced that the two-pronged approach is underinclusive. Once it is admitted that not *all* acts of a ‘crazy’ defendant are excused, the question how we should distinguish between acts that should be excused and acts that should not be excused has to be answered. I have tried to show that this can be done by establishing a legally relevant difference with the normal defendant. In M’Naghten’s case the difference with the ‘normal’ defendant – his inability to correct his delusion – is not legally relevant. Regardless of the unreasonableness of his belief, the facts as he presumed them to be would not amount to putative justification or ignorance of the law. Absent a legally relevant difference, it is hard to see what would merit a different treatment (i.e. acquittal where a ‘normal’ defendant would be convicted).

I am not sure that this analysis runs against moral intuitions, since it is grounded in the uncontroversial moral principle that to merit a different treatment, a rele-

46 Moore, ‘The Quest for a Responsible Responsibility Test,’ 662-63.

47 Cf. Morse & Hoffman, ‘The Uneasy Entente between Legal Insanity and Mens Rea,’ 1124.

vant difference has to exist. Even if it might seem counter-intuitive to some to hold M’Naghten responsible for his act, it would be in line with the legal principle that if a defendant was able to appreciate the wrongfulness of his act and control his behavior, he is responsible.

Of course, there may exist legally relevant differences that have been overlooked in the analysis above. Also, new scientific knowledge on mental disorders may shed another light on the (in)capacities of defendants with certain disorders. This, however, would not be fatal to the approach defended here. It would just mean that the prongs have to be reinterpreted or even that a new prong has to be added, to accommodate for these legally relevant difference.

## 11 Conclusion: the merits of the modern approaches to the insanity defense

The approach to insanity taken here demands for a detailed assessment of the psychological (in)capacities of the defendant at the time of the act, in order to establish to which legal standards the mentally disturbed defendant can be held. While the norms of the criminal law are the same for citizens that meet the criminal law’s assumptions of sound psychological functioning, the establishment of mental disorder makes it necessary to tailor the norms of the criminal law to the individual capacities of the mentally disturbed defendants.

The condition of a mental disorder of the two-pronged approach can therefore be viewed as legitimizing an inquiry into to what extent legal norms can be applied to the defendant. Both conditions (mental disorder and both prongs) thus are interdependent, the ‘excusing work’ cannot be done by either one of them alone.

A mental disorder does not necessarily mean that no normative standards can be applied at all. Some norms can be maintained (albeit sometimes less demanding ones) and some other standards just cannot be applied at all in a particular case. This nuanced, normative approach can only be codified in broad terms as is reflected by both prongs of the modern approaches to insanity. Generally speaking a disorder exculpates only if the defendant cannot appreciate the wrongfulness of his conduct or conform his conduct to the requirements of the law.

This approach acknowledges that there is something special about mental disorder. The regular, objective doctrines of putative justification, mistake of law or duress are not suited to do justice to mentally disordered defendants. Slobogin’s integrationist test does not account for legally relevant differences between ‘normal’ defendants and mentally disordered defendants.<sup>48</sup> The status approach that has attracted attention since the trial against Breivik and that has been theoretically elaborated by Moore should, however, be rejected as well. The latter places too much emphasis on the differences between mentally ill and ‘normal’ defendants.

48 See also Morse & Hoffman, ‘The Uneasy Entente between Legal Insanity and Mens Rea,’ 1123-32.

Johannes Bijlsma

Moore characterizes his approach as asking a judge or jury to decide whether or not the defendant is 'so crazy as to not be responsible.' This decision, as was explained above, is a commonsense judgment concerning the membership of the defendant of the moral community. As Moore acknowledges, similar phrased tests have been criticized as being too vague.<sup>49</sup> Putting the matter of vagueness aside, it is somewhat of a truism that a judge or jury must assess whether or not a defendant can be held responsible. Labelling a defendant as 'so crazy as to not be responsible' is not a satisfying explanation for an acquittal for (often) serious criminal acts.

Moreover, it is hard to think of a meaningful legal debate in the courtroom between defense and prosecution on whether or not this standard has been met. Under the interpretation of the two-pronged approach advocated here, the normative question would be: To what extent can this mentally disordered defendant be held to the same norms as other citizens? In the courtroom, defense and prosecution can debate to what extent the general legal norms should be tailored to the individual capacities of the defendant.

49 Moore, 'The Quest for a Responsible Responsibility Test,' 691.