EuJAP | Vol. 14 | No. 1 | 2018
UDK: 1:159.97
159.9:17

FALSE-POSITIVES IN PSYCHOPATHY ASSESSMENT: PROPOSING THEORY-DRIVEN EXCLUSION CRITERIA IN RESEARCH SAMPLING

Rasmus Rosenberg Larsen
University of Toronto at Mississauga

Original scientific article – Received: 17/06/2018 Accepted: 27/08/2018

ABSTRACT

Recent debates in psychopathy studies have articulated concerns about false-positives in assessment and research sampling. These are pressing concerns for research progress, since scientific quality depends on sample quality, that is, if we wish to study psychopathy we must be certain that the individuals we study are, in fact, psychopaths. Thus, if conventional assessment tools yield substantial false-positives, this would explain why central research is laden with discrepancies and nonreplicable findings. This paper draws on moral psychology in order to develop tentative theory-driven exclusion criteria applicable in research sampling. Implementing standardized procedures to discriminate between research participants has the potential to yield more homogenous and discrete samples, a vital prerequisite for research progress in etiology, epidemiology, and treatment strategies.

Keywords: *Psychopathy, PCL-R, False-Positives, Moral Psychology, Exclusion Criteria*

1. Introduction

Psychopathy is a personality disorder associated with interpersonal callous traits and antisocial behaviors (Hart and Cook 2012). The diagnosis is widely seen as one of the most researched and validated psychiatric disorders (Hare, Neumann, and Widiger 2012). Though the diagnosis has a history of diverse application, currently, it is primarily utilized in forensic psychiatry and psychology (Gacono 2016), presumably propelled by two orthodox beliefs about psychopaths, namely, that they are (1) disproportionately responsible for violent crimes (Baskin-Sommers et al. 2016; Reidy et al. 2015), which is intimately related to their (2) psychological incapacity to sufficiently grasp, and be motivated by, interpersonal moral values (Blair 2017; Hare and Neumann 2008; Stratton, Kiehl, and Hanlon 2015).

However, these orthodox views have shown to be particularly difficult to substantiate in scientific research. For instance, a large-scale meta-analysis by Yang and colleagues (2010) found no statistically significant relationship between diagnosed psychopathic *personality traits* and *violence* (similarly, see Kennealy et al. 2010; Singh, Grann, and Fazel 2011). Though the link between diagnosed psychopaths and general delinquency is moderate (e.g. Serin, Brown, and Wolf 2016), critics have argued that this is a trivial

epiphenomenon of data being collected from an incarcerated populace (Camp et al. 2013; Skeem and Cooke 2010a).

Next, the view that psychopaths have difficulty perceiving moral values, let alone be motivated by them, has been similarly difficult to prove. In a meta-analysis on diagnosed psychopaths' moral judgment and comprehension, Marshall and colleagues (2018) did not only find the hypothesis poorly corroborated, but also determined that there was "evidence against the view that psychopathic individuals possess a *pronounced and overarching* moral deficit" (Marshall, Watts, and Lilienfeld 2018, 47, original italics). Though a handful of studies suggest idiosyncrasies in terms of moral judgment in diagnosed psychopaths (e.g. Blair 2011; Glenn et al. 2009), these findings could hardly support the stronger claim that psychopaths are psychologically *abnormal* with regards to morality (for a similar conclusion, see Borg and Sinnott-Armstrong, 2013).

Some researchers have interpreted these (and other challenging) findings as altogether undermining for the research paradigm (e.g. Cavadino 1998; Jalava, Maraun, and Griffiths 2015; Mullen 2007), where others have proposed more cautious, auxiliary explanations. One such proposal is that the incongruities are not necessarily signs of a null-hypothesis, but instead signs of measurement error resulting from significant false-positives in research sampling (e.g. Cooke et al. 2004; Cooke, Michie, and Skeem 2007). Indeed, psychopathy studies are conducted on *diagnosed* psychopaths, meaning that samples are selected through applying conventional diagnostic criteria, which altogether casts a net that may be too inclusive, i.e. selecting individuals who are not *actual* psychopaths (false-positives) (e.g. Skeem & Cooke 2010a, 437). The consequence being that if our research sampling is inadequately representing of the diagnostic entity we purport to study, this is likely to show up in research results as tangible discrepancies.

One way to deal with these concerns is to use so-called *exclusion criteria* in research sampling. Exclusion criteria are a set of predefined conditions that are used to identify and exclude specific individuals from a study (Salkind, 2010). While exclusion criteria are common in psychopathy studies, it is typically used to minimize sample *contamination* (e.g. comorbidity, demographics, intelligence, etc.). In this sense, exclusion criteria are essential, but they are nevertheless poorly equipped to counter the more profound problem of false-positives: identifying individuals who meet the diagnostic threshold of psychopathy, but who are not *actual* psychopaths.

This paper aims to demonstrate how so-called *theory-driven* exclusion criteria are a viable method to counter the false-positive problem. Though this method is not novel to mental health studies, it has yet to be broadly embraced by psychopathy researchers. An example of a theory-driven exclusion criterion is "bereavement" in Major Depressive Disorder research, which selects for individuals suffering from depressive states due to losing a significant other (spouse, relative, etc.). Such practice signals a theoretical commitment, namely, that the depressive effect of losing a loved one is a *normal* reaction, and therefore not a valid sign of psychological *abnormality* (e.g. Wakefield and First 2012). The effect of such an exclusion criterion is that depression false-positives (e.g. people experiencing prolonged sadness of bereavement) are not included in the research, and therefore will not distort research findings. In a methodologically similar fashion, this paper draws on moral psychological research to develop tentative theory-driven exclusion criteria with the aim of locating psychopathy false-positives in sample selection.

The paper proceeds as follows: First, the relationship between theory, research paradigms, and clinical models is surveyed to highlight important complexities concerning false-positives. Second, one specific theory of psychopathy is outlined, which will serve as the benchmark for identifying false-positives. Third, on the basis of these false-positive examples a clear exclusion criterion is formulated. The paper is concluded with a brief discussion on how such an exclusion criterion can be implemented in existing data sets and research paradigms, leading to potentially more homogenous research sampling, better suited for proper scientific inquiry.

2. Structural Challenges in Psychopathy Studies

This section considers some general observations, challenges and structures of psychopathy studies, elucidating why false-positives have become an eminent problem. Though these considerations will appear mundane to some readers, they are nevertheless essential in explaining the general motivation behind developing theory-driven exclusion criteria, hereunder the potential advantages as well as challenges inherent to such a method.

Psychopathy studies are arguably as old as *modern* mental health research. When Benjamin Rush in 1786 suggested the existence of the disorder (which he called *anomia*), it was not only a novel contribution to the nosological nomenclature of his days, but also a proposal to expand the general concerns of the mental health profession to also include mood and personality disorders (Rush 1972). In Rush's day, mental illness was typically associated with either severe psychosis or cognitive underdevelopment, and abnormalities of mood, personality, etc., was believed to be separate from the scope of mental health (e.g. Goodey 2011). Rush's diagnostic label, then, pushed research interests in the direction of contemporary psychiatric concerns.

Larger professional concerns aside, Rush's newly proposed concept was more precisely an attempt to account for a specific patient phenomenon: those individuals who exhibited behaviors so obscene, irrational, and self-defeating, that they could hardly be seen as a variation of normal human conduct, but which could still not be explained with reference to insanity. Instead, Rush's explanation was that these individuals had a fundamental inability to distinguish between, and thereby be motivated by, the normative fabric of society. As a case illustration, Rush gave an anecdotal example of a young man named Servin who had a dazing intellect and pronounced social skills. However, Servin was also "treacherous, cruel, cowardly, deceitful, a liar, a cheat, a drunkard and a glutton, a sharper in play, immersed in every species of vice, a blasphemer, an atheist. In a word—in him might be found all vices that are contrary to nature" (Rush 1972, 7). With Rush's concept, medical men of the day now had a relatively simple and intelligible account of a complex phenomenon, namely, when seemingly *normal* individuals were behaving overtly antisocial: it was a disease of the moral faculty (see also Carlson and Simpson 1965).

Though the immediate time after Rush saw an excess of innovations of the psychopathy diagnosis (e.g. Sass and Felthous 2014), the central tenets of Rush's proposal are still, to this day, considered germane: psychopathic patients are not suffering from a disorder of the intellect, but instead, their disorder allegedly consists of an incapability to know right from wrong, premediating them towards antisocial conduct. This perspective was, for instance, dominant in Hervey Cleckley's opus, *The Mask of Sanity* (1988 [first edition in 1941]), a work that is broadly acknowledged as the bedrock of contemporary psychopathy research. As Cleckley observed:

> [The psychopath] is *unfamiliar* with the primary facts or data of what might be called personal values and is altogether incapable of understanding such matters. It is *impossible* for him to take even a slight interest in the tragedy or joy or the striving of humanity as presented in serious literature or art. He is also *indifferent* to all these matters in life itself. Beauty and ugliness, except in a very superficial sense, goodness, evil, love, horror, and humor have *no* actual meaning, *no* power to move him. He is, furthermore, *lacking* in the ability to see that others are moved. It is as though he were colorblind, despite his sharp intelligence, to this aspect of human existence. It *cannot* be explained to him because there is *nothing* in his orbit of awareness that can bridge the gap with comparison. He can repeat the words and say glibly that he understands, and there is *no* way for him to realize that he does *not* understand. (Cleckley 1988, 59, my emphasis)

While this general viewpoint—i.e. psychopathy as an inability to properly distinguish between right and wrong—has not been sufficiently supported by science (e.g. Borg and Sinnott-Armstrong, 2013; Lilienfeld 2018; Marshall, Watts, and Meffert et al. 2013; Marshall et al. 2017), its centrality must not be underestimated. For instance, if diagnosed psychopaths are not incapacitated in terms of moral psychology, it is difficult to see what makes them different from individuals who knowingly and deliberately behave immorally. Contrary to regular individuals who (for whatever reason) behave antisocially, psychopaths are hypothesized to stand out by having etiological mechanisms underpinning their antisocial behavior and attitudes. Thus, the term psychopathy was, and still is today, meant as a scientific hypothesis about a distinct class of people marked by a distinct psychology (i.e. a "personality disorder"); psychopathy was not proposed as a mere Latinized, non-scientific word for people who behave badly (in such a case, the expression "being a bad person" would suffice).

So, why has Rush's hypothesis proved so difficult to support in scientific research designs? One answer to this question is that the disorder does not exist; that there is no such thing as being moral-psychologically incapacitated. This simple answer is the null hypothesis of psychopathy research, and it is an answer that more and more researchers currently lean towards, perhaps evident in the attempts to reframe the scientific discussion (e.g. Blackburn 1988; Brzović, Jurjako, and Šustar 2017; Lilienfeld, Smith, and Watts 2016; Mullen 2007; Patrick 2006).

However, there are convincing auxiliary explanations for the discrepant findings. First, we must acknowledge the complexity of the alleged phenomenon before we draw precarious conclusions. Psychopathy studies are not dealing with concrete objects ready-made for scientific scrutiny. Instead, psychopathy studies begin with a rather elusive phenomenon, namely, a specific patient class characterized by an appearance of "normality", but who exhibit, all things considered, abnormal psychological symptoms (i.e. callous personality) and behavioral signs (i.e. antisocial). Quite literally, the research tradition starts with patients entering a clinical practice (e.g. via some kind of forensic institution), and as a result of their seeming abnormalities, researchers hypothesize that they make up a homogenous class, that is, a class of individuals sharing the same hypothesized medical condition. Thus, before we criticize the research, we must acknowledge that it is first of all excruciatingly difficult to accurately measure—based on signs and symptoms—which individuals belong to the alleged homogenous group of patients (e.g. Brazil et al. 2018). And if the phenomenon in and by itself is difficult to demarcate, we must also anticipate that the science about the phenomenon will be difficult to parse out.

A second related explanation for the discrepant findings is found in the general structures of how research on psychopathy is conducted. Psychopathy research can be divided (roughly) into three hierarchically interrelated domains: (1) theory formulation, (2) empirical test paradigms, and (3) clinical modelling. This structure mirrors general research efforts in mental illness (for a more detailed version, see Smith 2005):

(1) Theory Formulation: The first domain is a theory building effort, which essentially is an attempt to schematize intelligible research hypotheses that reflect—to the best of our knowledge—the psychological abnormalities (i.e. signs and symptoms) that are associated with the patient class. Whatever is formulated in this domain will be the basic guidelines for our research hypotheses and designs.

(2) Empirical test paradigms: The second domain is a data generating effort where various hypotheses are tested in different research designs. It is here that we figure out whether one hypothesis is better supported than others, or whether some hypotheses are falsified. Conclusively, the results in this domain will ideally feed back into the domain of (1) theory formulation, in an attempt to sophisticate these working theories.

(3) Clinical modelling: The third domain is where theories and research efforts come together with the aim of building models that can be utilized in clinical settings, for instance, in order to inform treatment and behavior prediction efforts. Psychopathy assessment tools are examples of the final product in this domain. This portrayal of the scientific field as a hierarchically interrelated domain-specific effort is, of course, a particularly idealized version of what actually goes on. Most research efforts are intimately involved in more than one of these domains at the same time, and often, in the same specific research design (e.g. Skeem and Cooke 2010b). Thus, when philosophers of science speak of such a division of research efforts, we must also understand that the practical reality is somewhat different. However, this does not remove from the fact that there exist discrete separations in terms of different scientific challenges (e.g. theory formulation, empirical testing of theory, and clinical application). But more importantly, it also serves to inform us in what ways crucial links are made between the different knowledge building efforts.

With this tripartite hierarchy in mind, we can thus locate the focus of the present contribution, i.e. the effect of false-positive *assessments* in research *sampling*. More specifically, how the clinical models of psychopathy can influence the dynamics within the two other research domains through a *backward relation*: when the clinical model (partly or wholly) functions as a potential source of measurement error in theoretical and empirical research efforts. This observation warrants a further explanation.

The lion's share of psychopathy research is conducted on *diagnosed* psychopaths, that is, individuals from forensic institutions. The most common method to diagnose patients with psychopathy is the *Hare Psychopathy Checklist-Revised* (PCL-R) (Hare 2003), which consists of 20 diagnostic items loading on two factors (dividable into four facets) of *interpersonal/affective* (Factor 1) and *social deviance* (Factor 2). Examples of the diagnostic items include *glibness/superficial charm*, *pathological lying*, *lack of remorse/guilt*, *shallow affect*, *poor behavioral control*, *impulsivity*, *irresponsibility*, *criminal versatility*, etc. A PCL-R assessment is provided by reviewing the patient record as well as administering a semi-structured interview. From this information, the patient is evaluated in accordance to the 20 diagnostic items, scoring each item from 0 to 2 points, where 0 is given if the item is not present, 1 if it is somewhat present, and 2 if the item is definitely present. As such, a patient can score on a scale ranging from 0 to

40, where a score of 30 is considered the *clinical* threshold for a proper diagnosis; though, *research designs* typically include patients scoring between 20-40.

Naturally, the quality of psychopathy research is therefore contingent on the ability of the PCL-R to discriminate non-psychopaths from psychopaths. That is, empirical research is dependent on the ability of the PCL-R to correctly select individuals who are *actual* psychopaths. Where the terminology "actual psychopaths" is here understood as the state where an individual meets the most valid *theoretical* criteria for psychopathy (i.e. level 1 in the hierarchy), which is not necessarily the same as meeting the *conventional* threshold for psychopathy (i.e. level 3 in the hierarchy). Indeed, it could be that the PCL-R is primarily build on one specific theory of psychopathy, which deviates from the theory we wish to test in a research design. One key challenge, then, is to develop a method that can accurately discriminate and exclude within the PCL-R sample, so the individuals we eventually select for a study are a more appropriate match for the theory of psychopathy we seek to test. This challenge is the topic of the following two sections.

3. Theory Formulation

The central proposal of this paper is to demonstrate how to identify theory-driven exclusion criteria, and how this can work in favor of reducing false-positives in research samples. Such a method begins by first detailing and formalizing an applied theory of psychopathy, and how to use such a theoretical formalization to develop exclusion criteria for research sampling procedures. As hinted at in the previous section, this method is based on the Popperian view of science, that our understanding of any phenomenon is a *theoretical understanding* (Popper 2002), meaning that a valid account of psychopathy is contingent on the ability of a theory to demarcate and predict features essential to the studied phenomenon. A theory-driven exclusion criterion, then, will rely mostly on specific predictions (i.e. necessary conditions) that follows from the relevant theory of psychopathy.

The first aspect that needs to be considered regards the choice of theory. Psychopathy studies are rife with unique theories offering everything from subtle to drastically different accounts of the clinical phenomenon (e.g. Blackburn 2006; Brazil and Cima 2015). Any attempt of developing theory-driven exclusion criteria will be wholly contingent on which exact theory is tested in the research design. For example, contemporary psychopathy research is mainly dominated by two competing types of theories, which we may refer to as: *Emotion Deficit Theories* (e.g. the Emotion-Based Learning Theory [Blair, Mitchell, and Blair 2005]) and *Cognitive Theories* (e.g. the Response Modulation Hypothesis [Hamilton and Newman 2018]). The former theories build on a view that psychopathy is an affective dispositional dysfunction (e.g. psychopaths cannot process certain emotions properly), and the latter posits psychopathy as a disorder of executive cognitive functions (e.g. psychopathy is underpinned by an attention deficit). Given the fundamental differences between these two types of theories, the exclusion criteria we derive will likewise be diverse. In other words, the view on whether psychopathy is an *affective deficit* or a *cognitive deficit* will also determine what exactly is rendered a false-positive in the class of PCL-R diagnosed individuals. Though the *method* proposed in this paper is generally applicable in any research paradigm, the remainder of the paper will focus only on *Emotion Deficit Theories*.[1]

1 It should be emphasized that these two theories are not necessarily mutually exclusive, that is, the signs and symptoms observed in the patient class could be caused by at least two distinct etiologies,

Emotion Deficit Theories have been formulated in several different versions and substantiated with different arguments (e.g. Blair, Mitchell, and Blair 2005; Cleckley 1988; Fowles and Dindo 2006; Hare 1998; Lykken 1995; McCord and McCord 1964), and for the current purpose it would be redundant to describe the differences between them. However, central to all versions is the universal claim that psychopathy is underpinned by a significant impairment of emotion processing, which consequently leads to impoverished moral learning, motivation, comprehension, etc. (e.g. Blair 2017). The stereotypical psychopath, according to these theories, is the calm and fearless person who cares for no one except (presumably) him/herself; the root of the careless, fearless personality being a deprivation of the type of emotional content, which most people take as a cornerstone of their life experiences. As the creator of the PCL-R, Robert Hare, so illustratively put it: "Completely lacking in conscience and in feelings for others, they [psychopaths] selfishly take what they want and do as they please, violating social norms and expectations without the slightest sense of guilt or regret" (Hare 1993, xi).[2]

Positing a link between emotional disaffection and moral psychological impairment conveys a specific philosophical commitment. Indeed, for this claim to make any philosophical sense, it must necessarily be the case that ordinary moral psychology (i.e. moral learning, values, motivation, judgment, etc.) is connected in some *non-trivial* way to our emotional perceptions. Put differently, if the claim is that the psychopathy disorder is one of emotional dysfunction, and this disorder is also causing moral psychological deficiencies, then it necessarily has to be the case that there is a *bona fide* psychological link between morality and emotions. If, one day, psychologists falsify such a hypothesis, the Emotion Deficit Theories of psychopathy will likewise cease to be meaningful.[3]

There are, however, good reasons to believe that the hypothesis is on stable ground. Recent developments in moral psychology have emphasized the salient role emotions play in moral judgments and beliefs (e.g. Bloom 2013; Greene 2013; Haidt 2012; Pizarro, Inbar, and Helion 2011; Prinz 2016). Though there are many different and sound theories of emotion-based morality (i.e. sentimentalism), one of the more recent and elaborate defenses of the necessary link between emotions and morality comes from Jesse Prinz's *The Emotional Construction of Morals* (2007). In this work, Prinz argues that moral beliefs are composed of emotional dispositions (i.e. sentiments), namely, that when we judge something to be moral or immoral, what is really going on is that we are associating either a positive or negative emotion with our experiences (Prinz 2007, 13-14). For instance, when a person utters the sentence, "it was immoral

---

which could amount to both affective and cognitive disorders (e.g. Moul, Killcross, and Dadds 2012). This possibility further underlies the importance of developing theory-driven exclusion criteria: i.e. theory-driven exclusion criteria will here be a viable method for testing multiple etiological hypotheses from the same base sample.

2 This stereotype is sometimes referred to as primary psychopathy, a term that was introduced by Ben Karpman (1941). There have been repeated concerns about the PCL-R also "tapping" into the notion of so-called secondary psychopathy, which is the neurotic, impulsive stereotype (e.g. Blackburn et al. 2008). This further underlines the importance of developing theory-driven exclusion criteria to separate primary from secondary psychopathy in the PCL-R (e.g. Skeem and Cooke 2010a).

3 Notice that this theory is not committed to the stronger claim that there is a link between emotions and normative ethics. Psychological theories of psychopathy are only committed to moral psychological statements, that is, the extent to which people perceive and relate to moral values.

when John stole Mary's bicycle", what is really reported is a *feeling of disapprobation* associated with the event of John stealing Mary's bicycle.

The stronger claim that Prinz is making is that moral beliefs are *necessarily* and *sufficiently* caused by emotional perceptions: moral claims *are* sentimental claims. Prinz argues that there are intuitive and empirical evidence of this. Many people would readily agree that their moral beliefs are visceral, that they have strong feelings associated with their moral values, or as Prinz puts it, that their moral judgments "ooze of sentiment" (2007, 13). We are taken aback, flabbergasted, or stunned when we witness a miscarriage of justice; and we are positively thrilled and grateful when we read about a local hero. However, while these mundane observations might convince you that emotions seemingly co-occur with morality, it could be that they are not causal or constructive parts of morality. It could be that morality is a fully rational process, which just happens to trigger our emotions in an intimate, reliable fashion.

Prinz (and like-minded theorists) rejects this criticism, emphasizing that it is difficult to square reason-based moral theory with empirical cases where moral judgments appear to be sufficiently composed of emotional content alone. For instance, in Haidt and colleagues' now famous study (2000), test subjects were asked to rationally qualify a variety of stereotypical moral beliefs (e.g. whether they believe it is immoral to eat discarded human tissue). The study found that when test-subjects were incapable of qualifying their moral beliefs (i.e. incapable of giving reasons for why they believe something to be immoral), the majority still chose to stick to their initial judgement, defending their moral position with unqualified statements of sentiment; an indication that emotions may play a sufficient role in moral evaluation (for similar findings, see Hindriks 2015; Rozin et al. 1999; Uhlmann and Zhu 2014).

There are other types of research that supports Prinz's theory, for instance, neurobiological studies that find moral task solving reliably activating the same brain regions that are also involved in emotional processing (Garrigan, Adlam, and Langdon 2016; Greene and Haidt 2002; Shenhav and Greene 2014). However, it is not the purpose of this paper to defend Prinz's theory. All that is needed for now is to acknowledge the theoretical link between moral emotional psychology and the deficits that psychopaths allegedly exhibit according to Emotion Deficit Theories. Prinz's work gives us a concrete theory for why a deprivation of emotional perceptions will cause a substantially different moral psychology. If an individual is deprived of emotional content, the theory predicts substantially different performances in moral psychological test-paradigms.

More importantly, however, what Prinz's theory allows us to do is to qualify our sample selection from a *contrapositive* argument. That is, because Prinz's theory states that morality is necessarily and sufficiently grounded in emotional perceptions, it thus follows theoretically that if a person has no (or severely deprived) emotional perceptions, such a person will then not be capable of having ordinary moral psychological states. Inadvertently, then, Prinz is giving a theoretical explanation for the clinical phenomenon of psychopathy research, i.e. individuals who are ignorant to the moral fabric of society are *necessarily* the individuals who are also deprived of emotional psychological states.[4]

4 Interestingly, Prinz discusses the phenomenon of psychopathy to a great extent, taking the existence of psychopaths as an empirically supporting element of his moral theory, namely, that his theory can predict moral-psychological abnormalities in people with affective deficits (Prinz 2007, 42-47).

Prinz's theory is also compatible with the medical definition of psychopathy as a *personality disorder*. Though mental health researchers in general tend to disagree on what exactly this term implies (e.g. Kendler, Zachar, and Craver 2011; Lilienfeld, Smith, and Watts 2013), it is uncontroversial to posit that the label "personality disorder" is not meant to indicate subtle abnormalities in personality. For example, just because a person has statistically outlying moral viewpoints, it does not follow that such a person is psychopathic. What is implied by the term "personality disorder" is the more profound notion of having *global abnormalities*, that is, psychopaths must have full-fledged problems with perceiving moral psychological content (as a derivative of them lacking emotional content). Similarly, when we speak of personality disorders defined by emotional instability (e.g. Borderline Personality Disorder [APA 2013, 663]), we do not infer this disorder from singular episodes or discrete, subtle instabilities. Instead, having a Borderline Personality Disorder means that the individual exemplifies emotional instability over a prolonged period, across related contexts and situations.[5] Thus, when we speak about psychopathy, the disorder is understood to be one of global impairment of emotion dispositions manifest across context and time, as opposed to minor subtleties over a few episodes. As we will see in the next section, this seemingly trivial observation becomes rather important when developing exclusion criteria.

4. Theory-Driven Exclusion Criteria

In the following section, one specific exclusion criterion is formulated. This is done by demarcating concrete instances of false-positives in a PCL-R assessment/sample, and, from this example, deriving necessary and sufficient conditions that capture such false-positives from a sample selection.

From Prinz's moral psychological theory, it may be tempting to conclude that psychopaths can be *spotted* by singling out individuals who appear to have no emotions associated with typical moral situations. We may therefore conclude that, in fact, many of the 20 items in the PCL-R are already doing exactly that (i.e. related to deprivation of moral emotions).[6] This approach builds on the aforementioned view that psychopaths are incapable of feeling for and with others, and therefore they will also stand out as those who are disaffected by ordinary morality.

While this viewpoint is not necessarily false, it shall be demonstrated that it misses a crucial aspect about moral psychology, namely, that there are many cases where people are perfectly capable of having emotions—and thereby capable of moral perceptions—but nevertheless fail to demonstrate this capability. To an outside observer (e.g. a psychologist), then, such cases will appear as proper instances of

5 Relatedly, the Diagnostic and Statistical Manual of Mental Disorders defines a personality disorder as an "enduring pattern of inner experience and behavior that deviates markedly from the expectations of the individual's culture, is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment" (APA 2013, 645).

6 It will be up for discussion exactly how many items of the PCL-R are substantially related to moral emotions. What seems to be obvious, however, is that the PCL-R naturally seeks to portray this central aspect of psychopathy. Arguably, the following 12 items are somewhat closely related to a lack of moral emotional perceptions: grandiose sense of self-worth, pathological lying, cunning/manipulative, lack of remorse or guilt, shallow affect, callous/lack of empathy, failure to accept responsibility for own actions, parasitic lifestyle, irresponsibility, juvenile delinquency, revocation of conditional release, and criminal versatility (Hare 2003).

psychopathy, and therefore count positively towards a diagnosis. However, since such cases are not examples of proper psychopathy—because the individuals in question are fully morally capable—these cases will then be false-positives. The following will account for three such examples of false-positives under the subtitles: (a) moral blocking, (b) moral dimensionality, and (c) moral plasticity.

(a) Moral Blocking: Cases of moral blocking refer to instances where a person holds a proper moral belief, say, believing that murder is wrong, but nevertheless fails to emote when experiencing or committing a homicidal act; an event that under normal circumstances would have triggered an aversive emotional reaction, and thereby resulted in perceiving the situation as properly immoral.

The actuality of moral blocking builds on a perhaps trivial observation about human capacities, namely, that we are particularly good at convincing ourselves of an alternative interpretation of reality; especially when it comes to the moral fabric of our everyday lives. As noted by Albert Bandura, "people do not ordinarily engage in harmful conducts until they have justified to themselves the morality of their actions" (Bandura 1999, 194). What Bandura is referring to is the peculiar cognitive ability to *overrule* or *reason out* one's own moral viewpoints. Examples of such cases are many. Sometimes these morally blocking reasons are developed on a socio-political scale. For instance, as noticed by Arne Johan Vetlesen, in the years leading up to the genocide of the Bosnian Muslim population in Serbia in 1995, the media was overtly engaged with rhetorical arguments portraying Serbians as the righteous and supreme people of the region, while vilifying and denigrating Bosnians as unvirtuous subhumans (Vetlesen 2005, 175-182). The argument here being, that even though we can assume that most Serbians (even in the time of civil conflict) believed murder to be morally wrong, there may have been psychological mechanisms at work, which would bar *some* people from experiencing genocidal acts as an immoral issue at all (e.g. Vetlesen 2005, 189-193).

This point is, of course, difficult to prove in a research setup, and concrete evidence of moral blocking is primarily anecdotal. A particularly illustrative anecdotal case of moral blocking we find in Daniel Goldhagen's book *Hitler's Willing Executioners* (1996). In this work, Goldhagen challenges a leading view that most evildoings under the Nazi regime were carried out by obedient minions, as opposed to maliciously willing individuals (e.g. Milgram 1974). In the following quote, Goldhagen contemplates what must have gone through the head of German police officers when they were ordered to grab, say, a captured Jewish girl and walk her into the woods where she was to be executed point blank. Is the policeman merely obeying authority; or is he more profoundly failing to perceive the moral fabric of the situation?

> It is highly likely that, back in Germany, these men had previously walked through woods with their own children by their sides, marching gaily and inquisitively along. With what thoughts and emotions did each of these men march, gazing sidelong at the form of, say, an eight- or twelve-year-old girl, who to the unideologized mind would have looked like any other girl? In these moments, each killer had a personalized, face-to-face relationship to his victim, to his little girl. Did he see a little girl, and ask himself why he was about to kill this little, delicate human being who, if seen as a little girl by him, would normally have received his compassion, protection, and nurturance? Or did he see a Jew, a young one, but a Jew nonetheless? Did he wonder incredulously what could possibly justify his blowing a vulnerable little girl's brains out? Or did he understand the reasonableness of the

> order, the necessity of nipping the believed-in Jewish blight in the bud? The "Jew-child," after all, was mother to the Jew. (Goldhagen 1996, 217-218)

No doubt that many of Nazi regime's executioners had deep moral quarrels (e.g. Browning 1992, 67-68), but it seems equally viable that some executioners were not perceiving, and thereby not feeling, the moral gravity of the situation. If the latter is the case, as Goldhagen would like us to believe, then there seems to be at least two explanations for such a phenomenon: perhaps the soldiers were fully incapable of perceiving moral values, namely, that they were full-blown psychopathic individuals; or perhaps some of the soldiers, while perfectly capable of perceiving moral values, simply did not perceive the situation as morally relevant. Speaking to the latter case, we can think of moral blocking as a *failure of perception*, as when a person is morally capable, yet his/her moral values are overruled, that is, *blocked* by other psychological mechanisms. To an outside observer, however, both individuals—the psychopathic and the morally blocking executioner—would appear as proper psychopathic, though, the latter case is a false-positive.

(b) Moral Dimensionality: Cases of moral dimensionality refer to situations where a person holds seemingly paradoxical moral views. For instance, in *situation x* a person believes that stealing is wrong, and in *situation y* the same person believes stealing is permissible. Thus, when such a person commits an act of theft in situation x or y respectively, he/she will have likewise diametrically different moral emotions associated with the act.

The reality of moral dimensionality speaks to a rather mundane observation about human nature, namely, that our moral beliefs are scarcely consistent (e.g. Hauser et al. 2007; Hooker and Little 2000). Though we arguably tend to speak about morality—and, in particular, our own moral values—as grounded in some sort of universal principle, the reality is that our moral landscape can be rather multifaceted. Consider, for instance, a typical case of moral hypocrisy: John comes home from a long day of hard work and is met by his neighbor who is furious that some boys vandalized his white picketed fence. John strongly feels with his neighbor that what the boys did was unacceptable. After they agree that something should be done about it, John walks into to the house and turns on the TV. He immediately starts laughing when he sees the news reporting from the local credit union, which has been vandalized by the same boys. Excited, John yells out to his wife: "It was about time that somebody stuck it to those bankers!"

In this case of hypocrisy, John holds two different sets of sentiments about the same basic action: John is disposed to feeling disapprobation when his neighbor's property is vandalized (i.e. feeling that this case of vandalism is wrong) and feeling approbation when the local credit union is vandalized (i.e. feeling that this case of vandalism is right). Examples of moral dimensionality can, however, be much more profound than frivolous hypocrisy. If we take a closer look at the anecdotal case from Goldhagen's work, we might be able to explain some of the *willing executioners* (i.e. the real moral evildoers of Goldhagen's popular book) as individuals who are merely carrying out what they believe to be moral acts. Like in the case of John's dimensionality, some of these soldiers might *feel* justified in ending the lives of people with Jewish heritage, while if they were asked to execute a person they saw as a moral equal, they would feel heavily against it.

Instead of holding universal beliefs about morality, the claim from moral dimensionality is that people appear to hold (at least some) beliefs that are nominalist in nature. In

Robert Lifton's book *The Nazi Doctors* (1986), he uses the term "doubling" to describe the phenomenon of Nazi doctors being capable of treating human lives differently depending on whether they practiced inside or outside the walls of the harrowing concentration camps. The doctors that worked on both sides of these walls could *divide their self* into one that was killing and one that was treating, without this seemingly paradoxical self-understanding showing manifest signs of discomfort (Lifton 1986, 421).

Lifton and Goldhagen's examples bring complex challenges to psychopathy assessment. As with moral blocking, the case of moral dimensionality raises clear instances that, to an outside observer, would appear as if we are dealing with proper psychopathic individuals, for example, the Nazi executioner or doctor who, disaffected and deprived of guilt, goes to *work* and participates in some of the most heinous crimes known to humankind. In these cases, however, we are not necessarily dealing with a psychopath, since when these men left the premises of Nazi operations, many of them presumably returned to their families as sincerely loving and caring fathers, the only seeming abnormality being the killings, which to us may appear psychopathic; but to them a dutiful moral act. Thus, in the case of childrearing, family men, we are necessarily dealing with a false-positive.

(c) Moral Plasticity: Cases of moral plasticity refer to situations where a person (all things equal) changes their moral belief (i.e. sentiment) about a specific issue over a course of time. For instance, when a person in *time 1* believes killing to be wrong, but in *time 2* now believes killing is justified. Over such a period, then, this person will have changed their emotional disposition from in *time 1*, feeling disapprobation about killing, to feeling approbation in *time 2*. While we evidently change many of our moral viewpoints throughout our lives (and thus, according to Prinz, we change our sentiments/feelings), what the example of moral plasticity also captures is how our emotional dispositions may change due to overexposure to specific conditions.

It is a rather uncontroversial observation that emotional dispositions are in and by themselves plastic entities. What this means is simply that we are likely to emote differently in a situation that occurs on a routinely repeated scale. As Elijah Millgram noticed, the first time we were confronted (probably as a child) with the sight of a street person in need of food and shelter, it is likely that we reacted with much more emotional vigor, compared to when we encounter such a person for the nth time (Millgram 1999). Similarly, we may speculate that the way experienced surgeons move the scalpel calmly through human flesh hardly resembles the first time they did it in Med School.

If emotions are plastic in this sense, it thus follows from Prinz's moral emotionist outline that our moral psychology is similarly plastic. It might turn out that the more we are exposed to morally troubling situations, the more attenuated our emotional dispositions become. Again, there are plenty of anecdotal evidence for such cases. Consider here an extreme case of plasticity taken from a letter correspondence between a Nazi soldier, Walter Mattner, and his fiancée, reporting from a systematic mass slaughtering of Jews in Belarus, which Mattner took an active role in:

> There's still something else I have to tell you. I was in fact also present at the enormous mass killings the day before yesterday. For the first truckload my hand trembled slightly when shooting, but one gets used to it. By the time the tenth truck arrived I was already aiming steadily and fired surely at the many women, children and infants. Bear in mind that I also have two babies at home, to whom these hordes would do the same, if not ten times

> worse. The death we gave them was a nice, short death, compared to the hellish torture meted out to thousands upon thousands in the dungeons of the GPU [Soviet state police]. (EHRI 2017)

It would perhaps be too simple to explain Mattner's (rather disturbing) moral beliefs as an instance of moral plasticity. However, we do see aspects of plasticity in the way he is emoting strongly to the mass murdering, but after "ten truckloads" his emotions are no longer present with the same vigor. On Prinz's theory, we could speculate that Mattner was, in fact, feeling aversion and thereby moral disapprobation throughout the initial killings, but after a while his views altered with his attenuated emotions. This is, of course, mere speculation. However, what the Mattner-example gives us, though, is a proper case of a potential false-positive in a PCL-R assessment. While Mattner's actions would most certainly meet the threshold for many PCL-R items, it is nevertheless possible that he has none of the affective deficits that Emotion Deficit Theories hypothesize. After all, Mattner is emoting strongly in the beginning of the killing.

The cases of moral *blocking*, *dimensionality*, and *plasticity* were presented here as a way of illustrating concrete cases of false-positives in a PCL-R assessment. Cases of such individuals would most likely be diagnosed with psychopathic traits (because of their behavior), but from a theoretical standpoint they do not meet the fundamental qualities of the disorder (i.e. the emotional deficit). If psychopathy is a true personality disorder, its signs and symptoms must be present over time and context. In neither of the cases thus described are the individuals exhibiting a disordered psychology, that is, a global impairment of the emotional and moral psychological capacities over and across context; their seeming deficit is an instantaneous peculiarity. Though their acts are disturbing and morally problematic, their state is not a valid sign of a personality disorder.

It must be expected that critical comments can challenge the reality, prevalence, and frequency of these case examples. Such critical comments are not only anticipated, they are also vitally needed. What has been discussed so far are mere initial sketches, where a detailed consideration is beyond the scope of the present paper. Having said this, one potential objection must be briefly considered.

Let us assume, for the sake of the argument, that moral blocking, dimensionality, and plasticity are true examples of how our moral psychology can function, and that these cases therefore are proper examples of false-positives. Nonetheless, one might still hesitate about their prevalence and frequency in psychopathy samples. Indeed, the discussed case examples were not necessarily everyday happenings, but ostensibly rather rare episodes of moral idiosyncrasies. How can such uniqueness be relevant to psychological practices?

There seems to be good arguments against such objection. For one, psychopaths are often described as being involved in rather obscene and unique antisocial behaviors, and the case examples thus described are not necessarily more unique on this already too bizarre scale. Secondly, the way the argument was presented in the different case examples might seem absurd, but they are readily applicable in the many moral demeanors we find diagnosed psychopaths to be involved in. For example, while we might look at the typical gangster as a cold-blooded psychopathic criminal, there are no reasons why the behavior of such a person could not (instead) be a case of, say, moral plasticity. We may speculate that in such a person's everyday life they have accepted a certain *moral standard* premediated by their dog-eat-dog environment. When such patients speak to the psychologist about, say, a homicidal conviction, the psychologist

might rate this episode in isolation, completely neglecting that, in fact, the hardened gangsters may hold strong feelings about friends and family, individuals they are readily prepared to give his life.

With these qualifying remarks in mind, we can thus derive some guidelines for a theory-driven exclusion criterion meant to capture instances of moral blocking, dimensionality, and plasticity. This criterion we may refer to as: *Sufficient Moral Psychological Emotions*. By detecting other moral episodes (across contexts) where the patient has demonstrated a capability of sufficient moral psychological mechanisms, namely, having feelings associated with moral values, judgments, beliefs, etc. researchers will be able to reliably discriminate cases of moral blocking, dimensionality, and plasticity from the class of PCL-R diagnosed individuals. Moreover, the exclusion criterion is meant to select individuals who are morally capable, from those who are incapable, which, according to Emotion Deficit Theories, is a consequence of having attenuated emotions. Inadvertently, this exclusion criterion is also consistent with the view that psychopathy is a personality disorder, meaning that psychopathy necessarily is manifest across contexts; thus, being capable of forming any sufficient moral psychological beliefs and judgements is a capability *necessarily* reserved for non-psychopaths.

If the argument thus far is reasonable, it may still be discussed whether the exclusion criteria should be introduced in sampling procedures as a taxometric criterion or as a dimensional criterion. In terms of the former usage, it would be a matter of detecting one instance (or a conventionally decided number of instances) of sufficient moral psychological judgments, beliefs, values, etc. In terms of the latter usage, it would be a matter of detecting to *what degree* a person can be said to exhibit the exclusion criterion (e.g. scoring it on a Likert scale from 1-7, as is sometimes the procedure with anxiety criterion in sampling [e.g. Zeier and Newman 2013]). Indeed, while we may be able to quantify the exact number of concrete cases of proper moralizing, Prinz's moral theory seems to imply that moral beliefs are not simple taxometric entities, but also dimensional in terms of the extent to which (i.e. emotional vigor) a person believes a situation to be morally relevant. The details of such considerations, however, go beyond the scope of this paper.

5. Concluding Remarks

In sum, this paper has suggested that one way to counter the false-positive problem in psychopathy research sampling is to introduce theory-driven exclusion criteria; and in relation to Emotion Deficit Theories, *Sufficient Moral Psychological* states were proposed to capture discrete false-positives portrayed under *moral blocking*, *dimensionality*, and *plasticity*. Importantly, this argument should not be understood as a "catch-all" for false-positives in PCL-R assessment/sampling, but relevant only for studies pertaining to Emotion Deficit Theories, and further, as a criterion for excluding specific false-positives within this theoretical outline. Thus, the current contribution is merely a methodological outline for formalizing and implementing one specific exclusion criterion for one specific theory.

This paper was motivated by recent concerns in the field about false-positives. For example, in their widely cited critique of the PCL-R, Jennifer Skeem and David Cooke emphasized that the assessment method was an inadequate mapping of the research domain, due to "underinclusion and overinclusion of people and of the construct [i.e. psychopathy] itself" (Skeem and Cooke 2010a, 436). The larger concern being that if the PCL-R continues to function as a *gold standard* for assessing and sampling psychopaths

for scientific research, our inquiries are bound to overlook the many subtleties articulated by various theories of psychopathy; subtleties that are essential for theory and construct validation (Skeem and Cooke 2010a, 437-439). Introducing *Sufficient Moral Psychological* states as an exclusion criterion is thus an attempt to accommodate specific demands in the field.

If this method is sound, however, it still remains to be implemented in future research in order to test its actualized positive and/or negative impact on data and analyses. While it should be relatively straightforward to apply such an exclusion criterion when composing research samples (e.g. during assessment procedures), the strength of exclusion criteria in general, and this criterion in particular, is that it can be implemented in existing data samples and as re-assessments of already concluded studies. In cases where data collection consists of detailed patient records, researchers will be able to exclude test subjects with a fair amount of discretion. However, in cases where data sets consist of the raw PCL-R item scores, researchers may look for scores on the following three items: lack of remorse or guilt, shallow affect, and callous/lack of empathy. If the subject does not score 2 points on these three items, it may be taken as an indication of reported emotional dispositions, which signals an inference to a (relatively) proper functioning moral emotional psychology.

Over and above the practical considerations, the implementation of exclusion criteria must (ideally) be guided by standardized definitions and procedures. Only through shared guidelines in sampling practices will it be possible to achieve the needed homogeneity in data samples, which is not only a prerequisite for meaningful data accumulation, but also for community wide and cross-disciplinary research efforts in terms of etiology, epidemiology, and treatment developments.

Acknowledgments

For comments and general encouragement, I wish to thank Lauren Schroeder. Thanks to the participants in my seminar, Forensic Psychopathology, at the University of Toronto Mississauga (spring, 2018) for stimulating and fruitful discussions. Also, many thanks to the Editors and Reviewers of *The European Journal of Analytic Philosophy*, who helped significantly improve the final version of this manuscript.

REFERENCES

APA. 2013. *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition (DSM-5). Washington, D.C: American Psychiatric Publishing.

Bandura, A. 1999. Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review* 3: 193-209.

Baskin-Sommers, A. R., C. S. Neumann, L. M. Cope, and K. A. Kiehl. 2016. Latent-variable modeling of brain gray-matter volume and psychopathy in incarcerated offenders. *Journal of Abnormal Psychology* 125: 811-817.

Blackburn, R., C. Logan, J. P. Donnelly, and S. J. D. Renwick. 2008. Identifying psychopathic subtypes: Combining an empirical personality classification of offenders with the Psychopathy Checklist-Revised. *Journal of Personality Disorders* 22: 604-622.

Blackburn, R. 1988. On moral judgements and personality disorders. The myth of psychopathic personality revisited. *British Journal of Psychiatry* 153: 505-512.

Blackburn, R. 2006. Other Theoretical Models of Psychopathy. In *Handbook of Psychopathy*, ed. C. Patrick, 35-57. New York: Guilford Press.

Blair, R. J. R. 2007. The amygdala and ventromedial prefrontal cortex in morality and psychopathy. *Trends in Cognitive Sciences* 11: 387-392.

Blair, R. J. R. 2011. Moral judgment and psychopathy. *Emotion Review* 3: 296-298.

Blair, R. J. R. 2017. Emotion-based learning systems and the development of morality. *Cognition* 167: 38-45.

Blair, R. J. R., D. Mitchell, and K. Blair. 2005. *The Psychopath: Emotion and the Brain*. Malden: Blackwell Publishing.

Bloom, P. 2013. Just Babies: The Origins of Good and Evil. New York: Crown.

Borg, J. S., and W. Sinnott-Armstrong. 2013. Do Psychopaths Make Moral Judgments? In *Handbook on Psychopathy and Law*, ed. K. A. Kiehl and W. Sinnott-Armstrong, 107-128. New York: Oxford University Press.

Brazil, I. A., and Maaike Cima. 2016. Contemporary Approaches to Psychopathy. In *The Handbook of Forensic Psychopathology and Treatment*, ed. Maaike Cima, 206-226. New York: Routledge.

Brazil, I., J. van Dongen, J. H. Maes, R. B. Mars, and A. Baskin-Sommers. 2018. Classification and treatment of antisocial individuals: From behavior to biocognition. *Neuroscience and Biobehavioral Reviews* 91: 259-277.

Browning, C. R. 1992. *Ordinary Men: Reserve Police Battalion 101 and the Final Solution in Poland*. New York: Harper Collins Publishers.

Brzović, Z., M. Jurjako, and P. Šustar. 2017. The kindness of psychopaths. *International Studies in the Philosophy of Science* 31: 189-211.

Camp, J. P., J. L. Skeem, K. Barchard, S. O. Lilienfeld, and N. G. Poythress. 2013. Psychopathic predators? Getting specific about the relation between psychopathy and violence. *Journal of Consulting and Clinical Psychology* 81: 467-480.

Carlson, E. T., and M. M. Simpson. 1965. Benjamin Rush's medical use of the moral faculty. *Bulletin of the History of Medicine* 39: 22-33.

Cavadino, M. 1998. Death to the psychopath. *The Journal of Forensic Psychiatry* 9: 5-8.

Cleckley, H. M. 1988. *The Mask of Sanity: An Attempt to Clarify Some Issues About the So-Called Psychopathic Personality*. 5th ed. St. Louis: Mosby.

Cooke, D. J., C. Michie, S. D. Hart, and D. A. Clark. 2004. Reconstructing psychopathy: Clarifying the significance of antisocial and socially deviant behavior in the diagnosis of psychopathic personality disorder. *Journal of Personality Disorders* 18: 337-357.

Cooke, D. J., C. Michie, and J. L. Skeem. 2007. Understanding the structure of the Psychopathy Checklist-Revised: An exploration of methodological confusion. *British Journal of Psychiatry* 190: 39-50.

EHRI. 2017. D05 Excerpts of letters from a police secretary in Mogilev to his wife. European Holocaust Research Infrastructure.

Fowles, D., and L. Dindo. 2006. A Dual-deficit model of psychopathy. In *Handbook of Psychopathy*, ed. C. Patrick, 14-34. New York: The Guildford Press.

Gacono, C. B., ed. 2016. *The Clinical and Forensic Assessment of Psychopathy: A Practitioner's Guide*. 2 ed. New York: Routledge.

Garrigan, B., A. L. R. Adlam, and P. E. Langdon. 2016. The neural correlates of moral decision-making: A systematic review and meta-analysis of moral evaluations and response decision judgements. *Brain and Cognition* 108: 88-97.

Glenn, A. L., R. Iyer, J. Graham, S. Koleva, and J. Haidt. 2009. Are all types of morality compromised in psychopathy? *Journal of Personality Disorder* 23: 384-398.

Goldhagen, D. 1996. *Hitler's Willing Executioners: Ordinary Germans and the Holocaust*. New York: Alfred A. Knopf.

Goodey, C. F. 2011. *A History of Intelligence and 'Intellectual Disability': The Shaping of Psychology in Early Modern Europe*. Burlington, VT: Ashgate.

Greene, J. 2013. *Moral Tribes: Emotion, Reason, and the Gap Between Us and Them*. Penguin Press.

Greene, J., and J. Haidt. 2002. How (and where) does moral judgment work? *Trends in Cognitive Sciences* 6: 517-523.

Haidt, J. 2012. *The Righteous Mind: Why Good People are Divided by Politics and Religion*. New York: Pantheon Books.

Haidt, J., F. Björklund, and S. Murphy. Moral dumbfounding: When intuition finds no reason. 2000. Manuscript.

Hamilton, R., and J. Newman. 2018. The response modulation hypothesis: Formulation, development, and implications for psychopathy. In *Handbook of Psychopathy*, ed. C. Patrick. New York: The Guilford Press.

Hare, R. D. 1993. *Without Conscience: The Disturbing World of the Psychopaths Among Us*. New York: The Guilford Press.

Hare, R. D. 1998. Psychopathy, affect, and behavior. In *Psychopathy: Theory, Research, and Implications For Society*, eds. D. J. Cooke, A. E. Forth and R. D. Hare. Dordrecht: Spinger, Kluwer.

Hare, R. D. 2003. *Hare Psychopathy Checklist–Revised*. 2 ed. Toronto: Multi Health System.

Hare, R. D., and C. S. Neumann. 2008. Psychopathy as a clinical and empirical construct. *Annual Review of Clinical Psychology* 4: 217-246.

Hare, R. D., C. S. Neumann, and T. A. Widiger. 2012. Psychopathy. In *The Oxford Handbook of Personality Disorder*, ed. T. A. Widiger, 478-504. New York: Oxford University Press.

Hart, S. D., and A. N. Cook. 2012. Current issues in the assessment and diagnosis of psychopathy (psychopathic personality disorder). *Neuropsychiatry* 2: 497-508.

Hauser, M., F. Cushman, L. Young, R. Kang-Xing Jin, and J. Mikhail. 2007. A dissociation between moral judgments and justifications. *Mind & Language* 22: 1-21.

Hindriks, F. 2015. How does reasoning (fail to) contribute to moral judgment? Dumbfounding and disengagement. *Ethical Theory and Moral Practice* 18: 237-250.

Hooker, B., and M. Olivia Little, eds. 2000. *Moral Particularism*. Oxford: Oxford University Press.

Jalava, J., M. Maraun, and S. Griffiths. 2015. *The Myth of the Born Criminal: Psychopathy, Neurobiology, and the Creation of the Modern Degenerate*. Toronto: University of Toronto Press.

Karpman, B. 1941. On the need of separating psychopathy into two distinct clinical types: The symptomatic and the idiopathic. *Journal of Criminal Psychopathology* 3: 112-137.

Kendler, K. S., P. Zachar, and C. Craver. 2011. What kinds of things are psychiatric disorders? *Psychological Medicine* 41: 1143-50.

Kennealy, P. J., J. L. Skeem, G. D. Walters, and J. Camp. 2010. Do core interpersonal and affective traits of PCL-R psychopathy interact with antisocial behavior and disinhibition to predict violence? *Psychological Assessment* 22: 569-580.

Lifton, R. Jay. 1986. *The Nazi Doctors: Medical Killing and the Psychology of Genocide*. New York: Basic Books.

Lilienfeld, S. O., S. F. Smith, and A. L. Watts. 2013. Issues in Diagnosis: Conceptual Issues and Controversies. In *Psychopathology: History, Diagnosis, and Empirical Foundation*, eds. W. Craighead, D. Miklowitz and L. Craighead, 1-35. Hoboken: John Wiley & Sons.

Lilienfeld, S. O., S. F. Smith, and A. L. Watts. 2016. Fearless dominance and its implications for psychopathy: Are the right stuff and the wrong stuff flip sides of the same coin? In *The Dark Side of Personality: Science and Practice in Social, Personality, and Clinical Psychology*, eds. V. Zeigler-Hill and D. Marcus, 65-86. Washington: American Psychological Association

Lykken, D. T. 1995. *The Antisocial Personalities*. Hillsdale: Lawrence Erlbaum Associates.

Marshall, J., A. L. Watts, and S. O. Lilienfeld. 2018. Do psychopathic individuals possess a misaligned moral compass? A meta-analytic examination of psychopathy's relations with moral judgment. *Personality Disorder* 9: 40-50.

Marshall, J., A. L. Watts, E. L. Frankel, and S. O. Lilienfeld. 2017. An examination of psychopathy's relationship with two indices of moral judgment. *Personality and Individual Differences* 113 (Supplement C): 240-245.

McCord, W., and J. McCord. 1964. *The Psychopath: An Essay on the Criminal Mind*. Princeton, NJ: Van Nordstrand.

Meffert, H., V. Gazzola, J. A. den Boer, A. A. Bartels, and C. Keysers. 2013. Reduced spontaneous but relatively normal deliberate vicarious representations in psychopathy. *Brain* 136 (Pt 8): 2550-2562.

Milgram, S. 1974. *Obedience to Authority: An Experimental View*. New York: Harper & Row.

Millgram, E. 1999. Moral values and secondary qualities. *American Philosophical Quarterly* 36: 253-255.

Moul, C., S. Killcross, and M. R. Dadds. 2012. A model of differential amygdala activation in psychopathy. *Psychological Review* 119: 789-806.

Mullen, P. E. 2007. On building arguments on shifting sands. *Philosophy, Psychiatry, & Psychology* 14: 143-147.

Patrick, C. J. 2006. Back to the Future: Cleckley as a Guide to the Next Generation of Psychopathy Research. In *Handbook of Psychopathy*, ed. C. J. Patrick, 605-617. New York: Guilford Press.

Pizarro, D., Y Inbar, and C. Helion. 2011. On disgust and moral judgment. *Emotion Review* 3: 267-268.

Popper, K. 2002. *The Logic of Scientific Discovery*. London: Routledge.

Prinz, J. 2007. *The Emotional Construction of Morals*. New York: Oxford University Press.

Prinz, J. 2016. Sentimentalism and the Moral Brain. In *Moral Brains*, ed. S. Matthew Liao, 45–73. New York: Oxford University Press.

Reidy, D. E., M. C. Kearns, S. Degue, S. O. Lilienfeld, G. Massetti, and K. A. Kiehl. 2015. Why psychopathy matters: Implications for public health and violence prevention. *Aggression and Violent Behavior* 24: 214-225.

Rozin, P., L. Lowery, S. Imada, and J. Haidt. 1999. The CAD triad hypothesis: A mapping between three moral emotions (contempt, anger, disgust) and three moral codes (community, autonomy, divinity). *Journal of Personality and Social Psychology* 76: 574-586.

Rush, B. 1972. *Two Essays on the Mind: An Enquiry Into the Influence of Physical Causes upon the Moral Faculty, and on the Influence of Physical Causes in Promoting an Increase of the Strength and Activity of the Intellectual Faculties of Man*. New York: Brunner/Mazel Publishers.

Salkind, N. J. (2010). *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE Publications Ltd.

Sass, H., and A. R. Felthous. 2014. The heterogeneous construct of psychopathy. In *Being Amoral: Psychopathy and Moral Incapacity*, ed. T. Schramme, 41-68. Cambridge, MA: MIT Press.

Serin, R., S. Brown, and A. Wolf. 2016. The Clinical Use of the Hare Psychopathy Checklist–Revised (PCL-R) in Contemporary Risk Assessment. In *The Clinical and Forensic Assessment of Psychopathy: A Practitioner's Guide*, ed. C. Gacono, 293-310. New York: Routledge.

Shenhav, A., and J. D. Greene. 2014. Integrative moral judgment: Dissociating the roles of the amygdala and ventromedial prefrontal cortex. *Journal of Neuroscience* 34: 4741-4749.

Singh, J. P., M. Grann, and S. Fazel. 2011. A comparative study of violence risk assessment tools: A systematic review and metaregression analysis of 68 studies involving 25,980 participants. *Clinical Psychological Review* 31: 499-513.

Skeem, J. L., and D. J. Cooke. 2010a. Is criminal behavior a central component of psychopathy? Conceptual directions for resolving the debate. *Psychological Assessment* 22: 433-445.

Skeem, J. L., and D. J. Cooke. 2010b. One measure does not a construct make: Directions toward reinvigorating psychopathy research – reply to Hare and Neumann (2010). *Psychological Assessment* 22: 455-459.

Smith, G. T. 2005. On construct validity: Issues of method and measurement. *Psychological Assessment* 17:396-408.

Stratton, J., K. A. Kiehl, and R. E. Hanlon. 2015. The neurobiology of psychopathy. *Psychiatric Annals* 45: 186-194.

Uhlmann, E. L., and L. Zhu. 2014. Acts, persons, and intuitions: Person-centered cues and gut reactions to harmless transgressions. *Social Psychological and Personality Science* 5: 279-285.

Vetlesen, A. J. 2005. *Evil and Human Agency: Understanding Collective Evildoing*. New York: Cambridge University Press.

Wakefield, J. C., and M. B. First. 2012. Validity of the bereavement exclusion to major depression: Does the empirical evidence support the proposal to eliminate the exclusion in DSM-5? *World Psychiatry* 11: 3-10.

Yang, M., S. C. Wong, and J. Coid. 2010. The efficacy of violence prediction: A meta-analytic comparison of nine risk assessment tools. *Psychological Bulletin* 136: 740-767.

Zeier, J. D., and J. P. Newman. 2013. Feature-based attention and conflict monitoring in criminal offenders: Interactive relations of psychopathy with anxiety and externalizing. *Journal of Abnormal Psychology* 122: 797-806.