Moralized Theories of Coercion: A Critical Analysis

John Lawrence Hill

Follow this and additional works at: https://digitalcommons.du.edu/dlr

Recommended Citation
MORALIZED THEORIES OF COERCION: A CRITICAL ANALYSIS

JOHN LAWRENCE HILL

There is an old story about a law professor, an appellate court judge and a trial judge who go pheasant hunting together. The law professor is the first to take aim at a bird he believes to be a pheasant. But he is not certain that it is, indeed, a pheasant. As he trains his sight on the creature, he begins to question whether his concept of pheasant-ness is really simply a false construction, a manifestation of some mode of false legitimation, and proceeds to deconstruct the concept of the pheasant to determine whether there really is such a thing as a pheasant for, if there is not, then this bird could not possibly be one and so . . . And, before the law professor has lingered long on his hyper-Cartesian self-indulgence, the bird takes wing.

Next, the appellate court judge spots the bird. As he takes aim, he too questions the nature of the creature. Unlike the law professor, however, the appellate judge has a far more pragmatic three-part test by which to determine whether this bird is, in fact, a pheasant. He first notes its color, then its wing markings, and finally its tail plumage, comparing these with the requisites of the same three-part test. Again, however, before he has completed the analysis, the bird flies off.

Finally, the trial judge spots a bird he believes may be a pheasant. Almost instinctively, and with a celerity that leaves his two companions in hushed embarrassment for his impetuosity, he takes aim and fires, Downing the bird instantly. Then, and only then, the trial judge asks himself, “I wonder if that was a pheasant?”

The tale of coercion is not unlike the story of the pheasant. While law professors (and philosophers) struggle to understand the nature of, and justification for, coercion, laymen (and trial court judges) appear to have a more or less cohesive notion of coercion and duress. As with so many other legal concepts, however, the terra firma of pre-analytic intuition appears to evaporate into a miasma of self-referential skepticism the closer we examine the nature and raison d’etre of coercion. Better to shoot from the hip than to go small game hunting for some Deeper Understanding.

In his book, Coercion,1 the political philosopher Alan Wertheimer undertakes the most exhaustive overview of the legal and philosophical terrain relevant to the issue of coercion of which I am aware. Roughly the first half of the book examines the contours of the concept of coercion as it has developed

---

* Associate Professor of Law, St. Thomas University. J.D., 1988, Ph.D. in Philosophy, 1989, Georgetown University.
in diverse substantive legal areas including contract law, criminal law, tort law and constitutional law, among others. Throughout this portion of the book, Wertheimer develops an account of the way in which the concept is employed in these various doctrinal areas. Then, in the second half of the book, he seeks to develop a philosophical defense for what he describes as a "moralized" theory of coercion.

While the following is an overview and critical analysis of Wertheimer's moralized theory of coercion, many of the conclusions will be relevant to moralized theories as a class. Furthermore, while Professor Wertheimer contrasts his "fully moralized" theory with partially moralized accounts, I will discuss the latter as a prelude to the former by way of comparing moralized accounts with the other prevailing theory of coercion, which views coercion as the product of an overborne will. In sum, I hope that the analysis here will be helpful in thinking about different species of moralized accounts of coercion generally.

Part I of this Article briefly lays out the taxonomy of various theories of coercion and then describes Wertheimer's theory. Part II of the piece analyzes and critiques moralized accounts generally, with particular emphasis upon Wertheimer's version of the moralized theory. Finally, Part III sketches the departure point for an alternative to both moralized theories and the more traditional accounts which explain coercion as a function of the overborne will.

**PART I**

There are two prevailing theories which seek to explain what coercion is and why victims of coercion are permitted a defense for their acts. First, there is what I will call the "traditional" theory of duress, which can be traced to the writings of Aristotle, and which received further explication at common law in Blackstone's Commentaries. The traditional theory views coercion as an overcoming of the will of the victim such that the resulting action is viewed as unfree, involuntary, or against one's will. Thus, defenders of the traditional account tend to conceptualize coercion as a type of excuse, rather than as a justification, because the underlying rationale for the defense is that the coerced actor is not responsible for her act. Though there is increasing skepti-

---

2. Id. at pt. 1, Law.
3. Id. at pt. 2, Philosophy.
5. William Blackstone, Commentaries on the Law of England, bk. IV, ch. 2 (1765). Blackstone summarized the rationale underlying all excuses, including duress, as "the want or defect of will." Id. at 21.
7. The distinction between excuse and justification defenses has been the subject of considerable scholarly analysis. The distinction has been cast in terms of the difference between condoning an act (justification) and holding that, while the act is otherwise morally or legally impermissible, the actor is not morally or legally responsible for its commission (excuse). See Kent Greenawalt, Distinguishing Justifications From Excuses, LAW & CONTEMP. PROBS., Summer
cism as to the validity of the traditional theory of coercion among philosophers and legal commentators, it remains the prevalent explicit rationale among contemporary judges in deciding actual cases.

In contrast to the traditional theory, moralized accounts of coercion maintain that legal claims predicated upon duress are at least partially a function of normative judgments about the nature of the situation and the right of the victim to respond in a certain way. In other words, a claim of duress is not simply a legal conclusion drawn from empirical premises concerning the psychological state of the actor—i.e., that the actor did not act voluntarily—as with the traditional theory. Rather, the determination that a particular case is coercive may flow from antecedent moral convictions that the putatively coerced actor possessed a kind of moral privilege to yield to the threat, or that no person should have to resist a similar threat. In sum, the defender of a moralized theory of coercion need not maintain that coerced acts are qualitatively different, with respect to the standpoint of the psychological state of the victim of coercion, from uncoerced acts. All that is absolutely necessary for the moralized account is to maintain that the victim’s predicament is morally different from that of the uncoerced actor—that it would not be fair to hold the victim of coercion responsible for his act irrespective of whether the coerced act is “voluntary.”

We can further distinguish between partially and radically moralized theories of coercion. Partially moralized theories are those that persist in utilizing some empirical criterion—typically the concept of voluntariness—to assess the coerciveness of a situation, but use some normative criterion in drawing lines that serve to distinguish the exculpatory from the non-exculpatory. Partially moralized theories tend to conceptualize voluntariness either as a hybrid, half-normative, half-empirical construct or, what might amount to the same thing, as an empirical state that occurs in varying quantities, depending upon the situation. Voluntariness might then be analogized to concepts such as quickness or health, which permit qualitative comparisons on the basis of facts (e.g., the speed of a process or the physical condition of a subject, respectively) but which employ a baseline normative criterion before a thing can be considered quick, healthy, or voluntary.

1986, at 89 (discussing the distinction critically).


9. See MODEL PENAL CODE § 2.09(1) (1962) (using the language of the traditional model, albeit with an objective standard). The Code states that “[i]t is an affirmative defense that the actor engaged in the conduct ... because he was coerced to do so by the use of, or a threat to use, unlawful force ... which a person of reasonable firmness in his situation would have been unable to resist.” Id.

The emphasis upon exculpatory, rather than justificatory, analysis is clear in contemporary case law. For example, one recurring issue is whether the threatened harm was imminent enough to warrant a defense. See, e.g., State v. Toscano, 378 A.2d 755 (N.J. 1977) (permitting the defense even where the threat is not imminent). This makes clear that modern courts continue to examine the effects of a coercive threat upon the actor, rather than a “lesser evils” weighing of the relative harms endemic of the threat and the offense, respectively.
A recent example of a partially moralized theory is found in the work of Professor Joshua Dressier, who argues that:

In short, duress is not like other excuses. The excusing basis is not merely empirical, but primarily normative. Unlike insanity, infancy and intoxication, the issue is not simply whether, as an empirically-verifiable matter, the actor lacked volitional or cognitive capacity... Duress excuses when the available choices are not only hard, but also unfair.\(^{10}\)

He goes on to argue that duress excuses when the defendant “lacked a fair opportunity to avoid acting unlawfully.”\(^{11}\) Fairness of opportunity is defined by reference to an objective standard: the question to be asked is whether, “in light of the nature of the demand and the expected repercussion from noncompliance, we could fairly expect a person of nonsaintly moral strength to resist the threat.”\(^{12}\)

It might be objected here that Dressler’s theory is entirely normative in that he utilizes an objective standard that takes no account of the subjective psychological state of the actor. Indeed, where an objective standard is interpreted as saying simply, “The law will not hold a person responsible for the choices made under a certain range of choice options because we think this would be unfair,” the standard would be completely moralized. This is because the prescriptive justification for the theory has to do with normative concerns regarding the fairness of the options rather than factual concerns about the volitional state of the actor.

For his part, Professor Dressler does not appear to go this far. He states that the exculpatory nature of duress is not “merely empirical,” implying that it is partially empirical.\(^{13}\) Moreover, the objective standard he proposes possesses an empirical component insofar as the test for coercion is measured by asking how much pressure the ordinary person—the person of nonsaintly moral strength—should be able to endure. To put it differently, while the move to an objective standard, here as in other legal contexts, is made for purposes of uniformity and because of inherent problems of proof in assessing the validity of subjective accounts of a person’s own mental states, the basis for the objective standard in Dressler’s theory is firmly rooted in psychological facts regarding the ability of the actor to resist. In sum, Dressler’s theory is partially empirical: according to his view, coercion still has something to do with the victim’s capacity to resist a particular threat. It is also, however, normative in that an objective standard is used to define the baseline by which coercive situations are distinguished from the non-coercive.

In contrast to partially moralized theories, radically moralized theories of coercion eschew any references to the voluntariness or freedom of particular acts, at least where voluntariness is defined as an empirical state reflecting the


\(^{11}\) *Id.*

\(^{12}\) *Id.* at 1367 (emphasis in original).

\(^{13}\) *Id.* at 1365.
actor’s psychological capacity to resist acting. They view coercion not as a function of the inner (moral-psychological) state of the actor, but as a normative assessment of the external options open to the actor—i.e., whether it is fair to hold a person responsible for an act given the range of alternatives available to him or her. As such, radically moralized theories focus upon such explicitly moral questions as whether the putative victim had a privilege to succumb to the alleged coercive proposal, whether the proposal itself was immoral, or whether persons under similar circumstances should be excused for acts committed under like conditions.

Wertheimer begins his analysis and defense of a radically moralized theory of coercion by examining the relationship between coercion and voluntariness. He draws a tripartite distinction between cases where there is an absence of volition, as in instances of physical compulsion, cases of defective volition, such as where the actor is insane, and cases of “constrained volition” characteristic of duress. While these are the three senses in which the term “involuntariness” is used in common parlance, Wertheimer argues that there are important moral differences among these cases. Most particularly, he suggests that the first two senses of the term represent empirical states of affairs reflecting the psychological condition of the actor. By contrast, involuntariness, in the sense of constrained volition, is basically normative in nature. Moreover, the first two senses constitute exculpatory conditions, while the constrained volition endemic to situations involving duress operates as a type of justification.

Wertheimer’s rejection of the traditional theory, which he calls the “moral will theory” of duress, begins with the classification of duress as a type of justification. Defenders of the traditional view of duress necessarily must conceptualize duress as a type of excuse: where duress is characterized as the overcoming of the victim’s will by the coercive stimulus, the rationale for duress is quintessentially exculpatory, rather than justificatory.

Wertheimer characterizes coercion as an agent-relative justification, thereby distinguishing duress from necessity, a justification that he views to be agent-neutral. This permits him to maintain that duress is a justification in the face of the most convincing argument to the contrary—namely, that duress is sometimes available in situations which are not normally viewed as justified. More specifically, acts are usually justified on the ground that the act

14. As we shall see, there are inherent ambiguities with respect to the use of the term “voluntary.” See infra notes 16-20 and accompanying text (discussing Wertheimer’s “rationalized” definition of voluntariness); see also infra Part III (discussing two other senses of the word “voluntary”).
15. WERTHEIMER, supra note 1, at 3-5.
16. Id. at 9.
17. Id. at 9-10; see id. at 46-48 (discussing the implications of void and voidable contracts).
18. Id. at 164.
19. Id. at 31.
20. Id. at 165-69.
21. Id. at 305.
22. See generally id. at 287-306 (discussing parallels of coercion in morality with coercion in the law).
23. Id. at 168.
committed produces less harm than what would have occurred in the absence of the act. Duress, however, is sometimes permitted even where the act leads to greater harm than what would have otherwise resulted.\textsuperscript{24} Wertheimer responds by arguing that the competing harms must be evaluated from the standpoint of the actor, rather than from a morally neutral vantage point.\textsuperscript{25} It is in this sense that he views duress as an agent-relative justification.

The heart of Wertheimer's theory consists of the observation that coercion claims require that a two-part test be satisfied. Throughout his exegesis of various areas of substantive law, Wertheimer argues that the same two-part analysis underlies the law's treatment of coercion, which he dubs the "proposal prong" and the "choice prong," respectively.\textsuperscript{26} The proposal prong requires that we ask whether the putatively coercive proposal, typically a threat, is immoral. The choice prong requires that we examine the choice of options with which the victim is faced to determine whether he should be held responsible for the resulting choice. Each of these prongs must be "moralized" in the sense that deciding whether the proposal is immoral and the choice is unfair are overtly normative assessments. The proposal and choice prongs are individually necessary and jointly sufficient to a determination that a particular act was coerced: coercion requires that the proposal is immoral \textit{and} that the choice is an unfair one to expect someone to make.

To take what is perhaps the paradigmatic example of coercion—the case of the gunman offering the pedestrian a "choice" between his money and his life—the analysis would permit a finding of coercion based upon the antecedent convictions that such a threat (the "proposal") is immoral and that the range of options from which the victim must choose (i.e., his money or his life) is not a fair set of alternatives. Other cases raise a plethora of more subtle questions which require normative answers. These include such hotly debated questions as whether offers, in contrast to threats, can ever be immoral in the sense that they may be said to be coercive.

The answers we give to questions such as these—and they are ultimately normative answers, he argues—will shape the contours of our definition of coercion. As Wertheimer notes, however, there is a sense in which our view of coercion must precede the answers we give to these questions. In other words, we must have reached a conclusion about whether a particular offer can be "coercive" before we can decide that the proposal prong can be satisfied by an offer. Thus, the radically moralized theory does not afford us a straightforward deductive proof for coercion claims without a kind of logical circularity. Wertheimer acknowledges this, but argues that there simply are no viable alternatives insofar as the traditional theory is unworkable.\textsuperscript{27}

\textsuperscript{24} One example of this is the recent abrogation of the "no murder" rule. At common law, duress was not permitted as a defense to murder. More recently, however, the Model Penal Code and a number of state decisions have permitted the defense in cases of murder. See Dressler, supra note 10, at 1370-74 (arguing that the recent elimination of this common law limitation in the Model Penal Code evinces the exculpatory nature of duress).

\textsuperscript{25} Wertheimer, supra note 1, at 168-69.

\textsuperscript{26} Id. at 30.

\textsuperscript{27} See id. at 8 (arguing that, if it is workable, an empirical theory would be more attrac-
Throughout the work, Wertheimer maintains that coercion claims are highly contextualized. In the most obvious sense of the term "contextual," the idea that coercion may vary from context to context is an unremarkable one. Where this is interpreted to mean simply that some contexts or situations present coercive conditions while others do not, the statement is obviously true. But Wertheimer means more than this. For example, he cites the example of a prisoner who confesses to a crime after the police threaten to beat him. He argues that the courts may decide that the threat was coercive and, thus, that the confession should be excluded from evidence, but that the prisoner's accomplices may not reach the same conclusion. He then asks whether both the courts and the accomplices can be correct, and answers:

Not if we assume that, barring some metaphorical uses of the term, the truth conditions of a coercion claim are always the same. But if, as I believe, the truth conditions of a coercion claim can vary with context, then the courts and the prisoner's accomplices may both be right.

This response raises fundamental questions about whether Wertheimer's theory is, in any sense, determinative or, to put it slightly differently, whether it provides any prescriptive guidance in the sense that it tells us when a given set of conditions is coercive and when it is not. Part of the difficulty stems from the ambiguous use of the term "contextual." If by saying that "coercion is contextual" we mean that each viewer may decide for himself whether a particular situation is coercive, this will have profound implications for whether the theory can be viewed to be determinative. We will have a good deal to say about this in Part II. For now, it is worth noting that moralized theories are not all necessarily "contextual," at least if this is what Wertheimer means by the term.

Finally, a word is in order concerning Wertheimer's theory of the relationship between coercion and voluntariness. It might be surprising that Wertheimer has—or thinks he needs—such a theory since the major thrust of his book is to argue in opposition to the traditional theory of coercion. In other words, since Wertheimer argues that coercion has little to do with the effect upon the will of the actor, a discussion of coercion and voluntariness might appear unnecessary at best, and might possibly be viewed to undermine his central thesis.

Wertheimer's response to all this is to argue that questions of voluntariness may indeed be relevant to the issue of coercion when voluntariness is itself construed as a fundamentally normative concept. Wertheimer explains the connection between voluntariness and coercion in the following manner: "[T]he argument will be that one acts voluntarily when one
acts (or should act) from certain motives or that one acts voluntarily when the factors that define one's choice situation stand in a certain relation to the principles that one does (or should) accept. More specifically, one acts voluntarily when she acts pursuant to a proposal and a range of alternative options that are consistent with principles that she does (or should) affirm. For example, a person cannot be coerced if the proposal with which he is confronted is permissible under the set of principles to which that person does (or should) subscribe (proposal prong). Similarly, he cannot be said to be coerced if the principles to which he adheres in defining the rules of distributive justice could generate the range of alternatives from which he has to choose in the putatively coercive situation (choice prong.)

The distinctly deontological cast to this argument holds that "coerced choices are not unwilled;" rather, they are "against one's will" in the sense that they violate one's system of considered values and judgments. Importantly, it is not relevant that the person does not like the proposal and alternatives with which he is confronted. As Wertheimer puts it, "reluctance and voluntariness can well go hand in hand." Instead, it is the fact that one is forced to act in contravention of one's principles that is central to the ascription of involuntariness.

It should be obvious that this use of the term "voluntary" does not comport with our everyday meaning of the concept. On Wertheimer's "rationalized" account, voluntariness becomes a function, not of our ability to resist a choice option, but of whether, under the system of moral principles to which we do (or should) adhere, we would hold a person responsible for the choice made in that situation. Again, when voluntariness is defined in this manner, coercion and voluntariness become normative proxies for one another. There persists that same normative circularity such that one cannot decide the voluntariness question without having pre-determined the coercion issue.

33. Wertheimer, supra note 1, at 301. Wertheimer writes:
On this view, B acts voluntarily when B succumbs to a proposal that A has a right to make, even if it is one which B finds unattractive and would prefer not to receive. Why? Because B himself is committed to the principles which grant A the right to make the proposal. On the other hand, B acts involuntarily when A makes an immoral proposal (a moral baseline threat) because A's proposal attempts to get B to act contrary to his deep preference that he not be made to act in response to immoral proposals. Id. (emphasis in original).

34. Wertheimer appears to waffle here regarding whether the defining principles are those which the victim does or should affirm. At one point, Wertheimer argues that it is the principles to which the victim is committed which count. Id. at 302. His discussion of "wantons" and egoists appears to confirm this. Id. at 303. On the other hand, he suggests that the defining principles are those which the victims should accept. Id. at 301. Both positions have problems. See supra notes 62-63 and accompanying text.

35. Wertheimer, supra note 1, at 302.

36. Id.

37. The theory of voluntariness is "rationalized" in the sense that it makes one's volitional capacity a function of reasons or, more accurately, considered moral judgments. Rationalized accounts can be distinguished from empiricist accounts of motivation, which hold generally that reason has no role in the motivational mechanism of human behavior. On the latter view, reason has, at most, only instrumental significance: it tells us how to satisfy some end. See Wright Neeley, Freedom and Desire, 83 PHIL. REV. 32 (1974) (discussing these competing conceptions of motivation); see also infra note 62 and accompanying text (discussing the rationalized account).
Wertheimer maintains that, though we may find this less satisfying than defining coercion in terms of some empirical concept of voluntariness, it is ultimately the only definition of voluntariness that makes coherent sense of the relationship between the two concepts.

PART II

The following sections will sketch in broad strokes the response to partially and radically moralized theories, which will be discussed in Sections A and B, respectively.

A. Partially Moralized Theories

Partially moralized theories attempt to bridge the gap between empirical facts about voluntariness and normative propositions concerning coercion by conceptualizing voluntariness as a commodity which occurs in varying degrees. This permits defenders of hybrid accounts, such as that of Professor Dressler, to characterize voluntariness and coercion as inversely proportional concepts which exist at opposite extremes of the same continuum. In this fashion, normative criteria are used to designate the line between acts that are voluntary and those that are coerced.

At first glance, partially moralized theories appear to combine the best of both worlds. They reflect the intuition that persons placed in coercive situations do act, in some sense, “against their will.” At the same time, partially moralized accounts appear to be flexible enough to serve the social and legal need for deterrence while simultaneously serving the ends of fair notice and uniformity by establishing a general standard applicable to all.

To the extent that partially moralized accounts are predicated upon questions of voluntariness, however, two types of problems arise. First, partially moralized theories must confront the same types of difficulties that bedevil the traditional theory of coercion. Second, the gradational view of voluntariness implicit in partially moralized accounts raises a number of additional issues.

The first problem stems from the fact that coerced choices are not “involuntary” in the usual sense of the term. In both legal and philosophical contexts, voluntary acts are defined simply as those in which the subject acts—i.e., where the subject chooses a course of action and then acts accordingly. They require no more than this. Coerced acts are entirely voluntary in

---

38. These are situations in which the actus reus requirement in the criminal law is not met, or where an act is not volitional for purposes of intentional torts. The Model Penal Code § 2.01(1) requires a “voluntary act,” and specifically excludes the following from the class of voluntary acts: reflex movements, bodily movements during unconsciousness, acts performed under hypnosis, and other acts not the product of a conscious determination by the actor. MODEL PENAL CODE § 201 (1962). Similarly, in tort law, a volitional act is one requiring a mental element representing some corresponding motivational force, such as an intention or desire, and a corresponding bodily movement. PROSSER AND KEETON ON TORTS 33 (5th ed., 1984). This dualistic test, requiring a physical and mental element, can be traced back at least as far as the writings of John Austin in the 19th Century, though it is evident in case law before then. See H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 97-98 (1968) (discussing Austin’s view of the act requirement).
this prevailing sense of the term. The fact that the individual may not want to perform any of the respective courses of action open to him is irrelevant to the question of voluntariness (though it might be relevant to excusing him for the act). Therefore, coerced acts cannot be distinguished from non-coerced acts on the ground that the former are involuntary.

To put it differently, the traditional theory of coercion depends upon our acceptance of the analogy between acts of physical compulsion—characteristic of "no act" situations—and acts precipitated by coercion (which are sometimes characterized as resulting from a kind of psychological compulsion, as when we say the "will is overcome"). Those who reject the traditional model view such talk to be a kind of figure of speech. They recognize that the impermissible limitation upon external options characteristic of coercion is distinct from issues of voluntariness, which go to a subject's capacity to carry out a chosen course of action. In sum, they distinguish external freedom—having differing options open to one—from internal freedom—being able to act voluntarily.

Independent of these issues is the problem raised by conceptualizing voluntariness as a gradational concept quantifiable by degrees. Very simply, the problem is that voluntariness does not admit of degrees. While sliding-scale accounts of freedom and voluntariness have attained some currency in common sense accounts of moral responsibility, and are even found enshrined in the law in such doctrines as the defense of diminished capacity, in fact such views of freedom or voluntariness are conceptually incoherent. The defender of the notion that voluntariness is an empirical commodity that occurs in varying degrees must tell us what this strange empirical commodity is. They must be able to point to some psychological fact about the actor that waxes and wanes with differing ranges of options such that a given act can be more or less voluntary. For reasons discussed earlier, it will not do simply to point to the obvious fact that the range of external options may be more or less palatable depending upon the situation. A limited or unwanted range of external options may render the actor's situation more sympathetic, even excusable, but this does not make the act less voluntary any more than a particularly palatable range of options would somehow render the choice made from among them more voluntary.

To the extent that any partially moralized theory links duress to the concept of voluntariness, as Professor Dressler's theory appears to do, it suffers from similar difficulties. On the other hand, perhaps such partially moralized accounts can be re-interpreted as justifying duress wholly in non-empirical

---

39. See MODEL PENAL CODE § 4.02(2) (1962) (providing for mitigation of punishment in capital cases where some mental defect has impaired the defendant's mental state or the capacity to conform his conduct). Other criminal law doctrines such as the heat of passion defense, more recently dubbed the "extreme mental or emotional disturbance" test, have served a similar function. The Model Penal Code provides that a murder charge may be reduced to manslaughter "when the murder is committed under extreme mental or emotional disturbance (EMED) for which there is reasonable explanation or excuse." MODEL PENAL CODE § 210.3(1)(b) (1962); see Joshua Dressler, Rethinking Heat of Passion: A Defense in Search of a Rationale, 73 J. CRIM. L. & CRIMINOLOGY 421, 466-67 (1982), for a limited defense of the EMED excuse.

40. For a more detailed analysis of the relationship between the availability of external choice options and voluntariness, see Hill, supra note 6.
terms. This leads us to a discussion of radically moralized theories of coercion.

B. Wertheimer’s Radically Moralized Theory

The following discussion will focus upon three central themes of Wertheimer’s philosophical defense of the radically moralized theory of coercion. First, I will examine his treatment of coercion as an agent-relative justification, rather than as an excuse. Second, and perhaps more centrally, I will explore Wertheimer’s contention that coercion claims are contextualized and the related question as to whether it is possible to put a truth value on coercion claims. Here, I will consider whether Wertheimer’s theory may be said to be determinative in any sense. Finally, I will survey his theory of voluntariness and its relationship to coercion.

1. Duress, Justification, and Excuse

While duress is generally treated as a kind of excuse in criminal law and contract law, legal theorists are fairly evenly split on whether to consider duress an excuse or a justification. Those who adhere to some form of the traditional theory of coercion tend to conceptualize duress as an excuse in that the coercive condition is viewed to overcome the will of the actor, thereby relieving the victim of responsibility for the act. Thus, theorists defending an exculpatory theory do so on the basis of the etiology of the coerced act. Conversely, those who view duress to be a kind of justification look to the consequences of the act. On this view, duress is viewed to be a kind of choice of evils justification similar to necessity. Duress can be analyzed this way because the law has placed substantive limitations on the types of consequences for which the defense will be permitted. Most generally, with certain exceptions to be mentioned shortly, a criminal defendant usually will not be allowed the defense where the act he committed in succumbing to the coercive threat is greater than the harm with which he was threatened in the event he did not comply. Similarly, the defense was not available at common law as a defense to murder. This seemingly consequentialist (justificatory) limitation was couched in exculpatory terms by the fiction that the will could (or should) not be overcome unless the threatened harm was roughly proportional to the harm created by the coerced act.

The rejection of the traditional theory of duress tends to blur the distinc-

42. Some commentators have argued that duress should be viewed as an excuse. See Dressler, supra note 10, at 1350; Fletcher, supra note 8, at 829-31; Michael S. Moore, Causation and the Excuses, 73 CAL. L. REV. 1091 (1985). Others defend the view that duress is a justification. LAFAVE & SCOTT, supra note 8, at 435; WERTHEIMER, supra note 1, at 166; WILLIAMS, supra note 8, at 755.
43. Dressler, supra note 10, at 1370.
44. See id. at 1373-74 (arguing that “[s]ociety also has a right to expect a person to demonstrate a higher level of moral strength when ordered to kill a hundred innocent children than when commanded to kill one”).
tion between justification and excuse. In other words, the distinction between
excuse and justification becomes considerably less significant in the absence
of the traditional theory. For example, on a moralized account, where a person
is permitted a defense simply because we have decided that it is not fair to
hold someone responsible for any act they have committed under a certain
constrained range of options, it becomes more difficult (and less important) to
decide whether the rationale is ultimately exculpatory or justificatory. It is
more difficult because the rationale can be characterized either in exculpatory
terms, as a defense predicated upon the idea that a person faced with a highly
constrained set of options simply is not responsible for their action, or in
justificatory terms, by arguing that net social utility is achieved by not holding
victims of coercion responsible. The distinction is also less important here
because, once the availability of duress is made to turn on the question as to
whether or not it is fair to hold someone responsible for a choice made under
a given set of conditions, issues of justification and excuse become secondary,
if not completely irrelevant.

In his account of coercion, Wertheimer attempts to capture both the excul-
patory and justificatory dimensions of coercion by classifying justification as
an agent-relative justification.45 Put simply, because he rejects the traditional
theory of coercion, he does not wish to conceptualize duress as an excuse. Yet
he must explain the fact that the defense is sometimes permitted where, on an
objective weighing of harms, the harm caused by the victim is equal to or
greater than the threatened harm in the event that he does not comply.46 For
example, imagine a case where the defendant-victim plants a bomb which
injures several people to avoid having his six year-old daughter sexually as-
saulted by the kidnapper-coercers. The agent-relative qualification permits
Wertheimer to maintain that, while an objective weighing of harms might
preclude the defense here, the law must take special account of what is argu-
ably the natural human sentiment to wish to protect one's own daughter, even
at the cost of a greater risk to many others.

While we find cases such as this intuitively sympathetic, I would like to
suggest that there are fundamental difficulties with characterizing the reason
for this as justificatory in nature. The problem is that the agent-relative qualifi-
cation permits us to privilege certain aspects of human personality, values,
beliefs, and desires, etc., in a manner that ultimately undermines the notion of
justification. The point is that justification defenses are essentially "objective"
in the sense that society assigns differing values to two competing courses of
conduct, permitting a privilege for conduct that, in other circumstances, would
be impermissible. The more we take into account the particular individual's
personal, even idiosyncratic, beliefs, values, and feelings the more the defense

45. Id. at 168.
46. Strictly speaking, where the rationale for duress is viewed to be exculpatory such that the
victim is not responsible for his act, the gravity of the resulting act should be irrelevant. While
criminal law does place substantive limitations on this by requiring a proportionality between
the threatened harm and the act, it nevertheless permits a defense in situations where justification
could not apply, such as where the harm created by the victim is greater than or equal to the
threatened harm.
begins to look like an excuse. The man who saves his daughter from sexual assault at the cost of greater risk to others is not really justified in performing his act. Rather, we feel sorry for him. We empathize. And we decide accordingly to hold him not responsible for the act.

Nor can Wertheimer defend the agent-relative rationale against this objection by arguing that only "rational" agent-relative decisions will be justified. The introduction of rationality is merely another way of privileging some aspects of personality while excluding others under the guise that some are rational. Put differently, on one (subjective) interpretation of rationality, avoidance of even the most minimal harm to self would be rationally required, even where this entails the exposure of others to great risk. If, on the other hand, by "rational" we mean to impose some objective, agent-neutral analysis concerning the proper distribution of harms, then we have relinquished the agent-relative account altogether. We simply cannot have it both ways. We cannot have an agent-relative account that nevertheless incorporates within its justification some objective standard regarding what a person should do in certain circumstances.

Nor is there a middle way between the two extremes—e.g., by contending that an agent-relative rationale permits some departure from the permissible range of outcomes generated by an agent-neutral justification, while precluding some outcomes as illegitimate, beyond the pale of the "rational." To take, once again, the example of the father who is coerced to save his young daughter, suppose he chooses to save her at the cost of wiping out all civilization on earth. Would this be justifiable on an agent-relative account?

Wertheimer faces a dilemma here: if he concludes that such a decision might be rational, then we are back to the slippery slope. If this is rational, anything might be justified as rational, in which case anything could be justified. In this event, the category of the justificatory collapses into the exculpatory. On the other hand, if he argues that it is not rational, then why not? Why couldn't the rational man who is particularly devoted to his child choose her life or well-being to the continued existence of our entire civilization? In short, if this is not rational, then what is? Who gets to define what counts as rational and what exactly could the criteria be? Finally, even if we could specify some criterion for agent-relative rationality, how could any departure from an agent-neutral justification remain rational without giving up the "objectivizing function" of rationality?

Misconceiving the nature of duress will have profound implications for the way in which the defense is viewed, and for the cases to which it applies, even with the sometimes fuzzy distinction between justification and excuse. Duress should be viewed as inherently exculpatory in nature. Adopting an exculpatory theory of duress, however, does not require that we embrace the traditional theory of coercion, as we will see in Part III.

2. Contextuality and the Determinateness of Duress

We now turn to the epistemological core of Wertheimer's theory of coercion. He elaborates upon a theory of coercion which he characterizes as both "moralized" and "contextual." The use of these terms highlights the central tension in his theory between the theory having some prescriptive content and his claim that duress is context-relative. To put it starkly, it simply is not clear whether Wertheimer believes one can ever put a truth value on coercion claims such that they are determinative. Let us briefly survey his general theory to see how this might be a difficulty.

As we discussed in the previous Part, according to Wertheimer, coercion consists of a two-pronged analysis which requires us to examine both the nature of the proposal (e.g., whether it is a threat or an offer) that is the subject of the coercion claim, and the range of choices open to the putative victim. For a situation to be considered coercive, it must involve an immoral proposal to which the victim is morally or legally privileged to succumb. What constitutes an "immoral proposal" and under what circumstances is the victim so entitled? Wertheimer's answer is that these are highly contextualized issues; there is no set of necessary and sufficient conditions which is determinate in specifying when coercion exists. Rather, issues which arise pursuant to either the proposal prong or the choice prong must be "moralized"—they must be evaluated from the standpoint of a moral baseline.

For example, to know whether the proposal is "immoral," we must ask normative questions about such matters as whether the proposal adversely affects the ex ante options of the putative victim (e.g., is the "proposal" a threat that reduces her options?), whether it is substantively immoral (i.e., does it propose an immoral exchange?) and whether it exploits the victim in virtue of her background conditions. Similarly, the choice prong requires that we make normative judgments regarding whether the range of choices open to the putative victim were sufficiently limited to preclude moral and legal liability.

A central question must now be addressed: Does having a moralized or contextualized account of coercion entail that there really are no truth conditions for coercion claims—i.e., that we cannot assign a truth value to the questions raised pursuant to the two prongs of Wertheimer's theory. When we say that a particular person was coerced in a given situation, are we saying anything about the world at all, or are coercion claims simply subjective labels attached to a situation after the speaker has decided that he or she should wish to see the victim released from legal responsibility?

If coercion claims are not truth functional in the way that empirical statements are, then a theory of coercion cannot be determinate and particular cases of coercion will not be determinable. Most basically, in this event, the theory will not be able to tell us when coercion exists and when it does not. It will not be able to settle differences when there is disagreement concerning particular situations. Further, it will not be able to generate a set of outcomes

48. Wertheimer, supra note 1, at 30-46 (discussing the sorts of considerations relevant to coercion on a moralized theory).
which can be compared with existing law by which we could gauge the accuracy of the theory (or the law.) Similarly, because coercion claims remain indeterminate, there is nothing to measure against our normative judgments, and consequently no way of judging whether the theory comports with our pre-analytic moral intuitions in any particular case. Moreover, a theory that cannot tell us when coercion claims are satisfied will be unlikely to tell us why we permit a defense for coercion as an initial matter. It can only generate the conditional conclusion that if we view a certain choice situation as sufficiently unfair and if we view a certain proposal as immoral, we will have made out a claim of coercion.

Given these considerations, is Wertheimer's moralized account indeterminate? The answer is unclear, and it is here that the central tension in his theory lies. Wertheimer maintains that a coercion claim with a given descriptive or normative force will have certain correlative truth conditions. Roughly speaking, the truth conditions of a coercion claim are what must be the case for the coercion claim to be valid or acceptable. He goes on in the next sentence, however, to say that he does "not want to put much weight on the term 'truth,'" and further states that the truth conditions will sometimes depend upon empirical and sometimes upon normative premises. While this may indicate some provisional commitment to a determinate theory of coercion, this assumption is dramatically undercut in his discussion of "coercion and contextualism." It is in this discussion that he presents the prisoner example mentioned previously, where he states that two persons with contradictory judgments of the same situation can be equally correct about whether a given situation is coercive.

By maintaining that the truth conditions of a coercion claim may vary with context, Wertheimer might mean any of a number of different things. He cannot simply mean that, as the context changes, so too do our intuitions about whether the situation is coercive. This would simply state the trivial truth about which no one would argue: some situations are coercive while others are not. Rather, he states that the truth conditions for a coercion claim vary with context. This seems to indicate not simply that our conclusions vary from context to context, but that the conditions for our conclusion (or our application of the concept) might vary from context to context — i.e., that more might be required for a claim of coercion in one context vis-a-vis another.

Wertheimer gives as an example of this contextual variability the way in which we use the term "cold." The person who believes that a thirty-five degree January day in Vermont is "cold" will be wrong insofar as the temperature is actually mild, "given the context." But if the same conclusion is reached by someone newly arrived from Miami, they would be correct. Thus, the argument runs, contradictory opinions can both be correct.

The problem with this response is that the term "cold" here is used in two
distinct senses. The first sense of the term "cold" is used in a statistical sense referring to the normal range of temperature in Vermont during January. It is false to say thirty-five degrees is cold in this sense, whether the remark is made by the native of Vermont or the newly arrived Floridian. The second sense of the term "cold," on the other hand, is used to reflect the speaker's subjective impression regarding the temperature. In this sense, the Floridian's statement that it is cold is true, just as would be a similar statement by a native of Vermont if he simply happened to feel cold. Thus, Wertheimer's example commits the fallacy of equivocation in the very discussion in which he warns against the possibility of equivocation.53

Is there any other sense, then, in which the truth conditions of a coercion claim vary from context to context while rigidly designating the same concept of coercion from one context to the next? Perhaps we might think the account could be salvaged by asserting that the same fact will take on different significance from one context to the next. For example, thirty-five degrees will be statistically cold in Miami in January while not being so in Vermont. In this sense, it might be argued, the truth conditions for coldness (or coercion) will vary from context to context.

This move, however, is of little use in defending the contextualized account of coercion because, in this example, we have not changed the truth conditions for the application of the concept "cold;" we have merely picked out one fact (i.e., the fact that it is thirty-five degrees) and recognized that this is cold in one context and not cold in another. It is not that the truth conditions for coldness have changed; rather, the truth conditions for coldness are always measured relative to some background conditions (the geographical locale and time of year) which themselves vary. In sum, the concept of coldness is univocal though its application is relative to the context being evaluated.54 The problem with Wertheimer's account is that, while the truth value of some propositions (e.g., thirty-five degrees is cold) are context-relative, the truth conditions for any such proposition cannot vary without varying the meaning of the concept itself. Maintaining that the truth conditions of coercion vary with context is like saying that coldness will be defined as any temperature that falls below the average temperature for a region in one context, while stipulating that coldness in another context will mean any temperature in the bottom ten percent of the normal range of temperatures for that region. Varying the truth conditions for a concept means abandoning any univocal concept. It is simply a logical contradiction to claim that one is talking about the same concept from one context to the next, though one is applying different truth conditions to the different contexts.

The only other possible route for Wertheimer's theory is to assert that, by claiming that the truth conditions for coercion vary with the context, he means

53. Id. at 182.
54. Thus, if we operationally define as "cold" any temperature that falls below the average temperature for a particular region at a particular time of the year, the truth conditions for coldness will be the same in Miami as they are in Vermont. Although, of course, the background conditions by which coldness is measured will vary from one region to another.
that the concept of coercion is dependent upon the perspective of the viewer—i.e., that one person may legitimately view a situation as coercive while another may not, where both are correct. His example of the prisoner and his accomplices seems to bear out this interpretation insofar as the courts and the prisoner’s accomplices could both reach accurate contradictory conclusions. But if this is the correct interpretation of his theory, then he has given up any semblance of a determinative theory of coercion. Moreover, where the truth of coercion claims are simply a function of the subjective beliefs of the viewer, it is simply facetious to speak of there being “truth conditions” for coercion claims at all. We may, of course, assign a truth value to the subjective beliefs of differing individuals—i.e., it may be true that one person feels that a certain situation is coercive while another does not. If this is the move, however, then the truth conditions do not even have anything to do with coercion anymore; they are merely a function of what any individual subjectively believes. In sum, if what it means for the truth value of a coercion claim to be “contextual” is that different people feel differently concerning what constitutes coercion, then Wertheimer has told us something trivially true about persons and nothing at all about coercion itself.

In the end, it appears Wertheimer wants to have his truth and eat it too. He wants to be able to say that there are truth conditions for coercion claims because, if there are not, there would be no point in having a theory (indeed, if we cannot put a truth value on the conditions for duress, in what sense can they be “conditions” for duress?). At the same time, however, he seeks the perhaps infinite flexibility of a protean conception of coercion that would permit two persons to reach contradictory conclusions about the coerciveness of the same identical situation where both are correct in their assessments. If this is what it means for truth conditions to be contextual, then the idea of “truth” itself has lost all meaning.

Having concluded an examination of this aspect of Wertheimer’s theory, it should be noted that it is still possible to have a “moralized” and “contextualized” theory of coercion without relinquishing determinacy. The problem with Wertheimer’s account is that it employs relativized versions of these two concepts in a way that renders the theory hopelessly indeterminate. First, coercion claims can be context-relative in the sense that particular facts will take on different relevance in varying contexts. Just as thirty-five degrees is “cold” in one context and not in another, so too, given certain background conditions, a particular proposal might be coercive in one context and not in another. This does not mean, however, that the existence of coercion depends upon the subjective impression of the viewer.

Nor does a moralized theory of coercion entail that coercion claims are completely relativized. A moralized theory may be truth functional and fully determinate given one’s adherence to some form of moral realism, as opposed to some species of moral relativism or moral noncognitivism.55 Professor

55. “Moral realism” is used to refer to any view that holds that moral propositions reflect some external reality, or are true or false in something like the way empirical statements are. “Moral relativism” is a meta-ethical view that rejects moral realism—i.e., by arguing that moral
Wertheimer’s moralized theory of coercion is not truth functionally determinative because it does not take seriously the possibility of moral realism, as his prisoner example makes clear. Indeed, perhaps the deepest irony of Wertheimer’s account is that he seeks to provide a “moralized” account of coercion that cannot tell us what the normative parameters of coercion are.

Part III of this article sketches the parameters of a moral realist theory of coercion that avoid the pitfalls of relativized accounts such as those considered here.

3. Coercion and Voluntariness

Over the course of the past twenty-five years, a number of philosophers have offered differing versions of a novel way of understanding the relationship between motives and values in a way that permits a distinction to be drawn between voluntary acts and free acts. On such accounts, not every voluntary act is free. To put it differently, some sources of human motivation might meet the technical requisites of the “act” requirement (such that they are considered voluntary) without constituting an act which we would consider to be “free.” Perhaps the paradigmatic example of this is the heroin addict, whose use of the drug is voluntary in the technical sense, even as he wishes to be free of the addiction.

In a seminal article, Harry Frankfurt argues that, for an act to be free, it is not sufficient that the act simply flow from a first-order mental state such as a desire. Acting from a first order mental state may characterize voluntary behavior, but animals act voluntarily in this sense, as even Aristotle pointed out. What distinguishes persons from animals and free choice from mere voluntariness is the ability to form second-order mental states—i.e., mental states about other mental states. One acts truly freely, on this view, only when he acts from a first-order desire which is itself connected with, or the product of, a second-order value, desire or judgment. In other words, genuine autonomy comes in valuing (in a second order way) that which one desires as a first order matter. We rightly characterize as unfree the addict and the neurotic-compulsive because they act voluntarily in response to (first order) desires that are in conflict with second order desires. In short, the addict is unfree because, though he desires heroin, he does not desire that he desire heroin.

With a different twist, Gary Watson has argued that genuine free agency propositions reflect the individual’s (individual relativism) or his culture’s (cultural relativism) view of what is right and wrong. “Noncognitivism” is the view that we cannot know the truth or falsity of moral propositions—i.e., that moral statements do not refer to some knowable object.

56. Id. at 181. There simply is no way to reconcile a moral realist account of coercion with the claim that two persons can reach diametrically opposed positions about the same situation and both can be correct.


58. Aristotle, supra note 4, at 935, 938.

59. An example of a second-order mental state is having a desire to inculcate a first-order desire to enjoy classical music, or to have a (second-order) intention to restructure one’s first order desires. In short, second order mental states are mental states that have as their intentional object some first-order mental state.
exists only when one acts from motives that are themselves in accord with our 
"evaluational system." In sum, our acts must flow from voluntary impulses 
which are themselves the product of our considered values and judgments. 
This evaluational system serves not only to ground our actions in our deepest 
principles and values, but also rationalizes our values and beliefs, making 
them consistent with one another.61

Wertheimer's discussion of voluntariness follows the tradition of such 
rationalized accounts of volition. He distinguishes the "volitional" from the 
"voluntary," using the term "volitional" to designate the traditional sense of 
the term "voluntary," as where the person has acted in the technical legal 
sense.62 In contrast, Wertheimer argues that an act is not voluntary, for pur-
poses of duress, if it is performed as the result of an immoral proposal that 
leaves the actor with a morally unacceptable range of choice options. These 
two prongs are "moralized" in accordance with a set of moral principles to 
which we do (or should) adhere. A range of choices that fails to conform to 
these principles, Wertheimer contends, are involuntary in the sense that they 
are "against one's will," even if they are otherwise "volitional."

There are three distinct difficulties with this approach to voluntariness, at 
least as it relates to duress. First, even if voluntariness were a function of 
some set of moral principles—even if it could be rationalized in the fashion he 
suggests—Wertheimer never resolves the inherent ambiguity as to whether 
voluntariness is defined by reference to the moral principles to which we do or 
those which we should conform. Second, regardless of his answer to the first 
problem, the voluntariness of an act is simply not a function of the system of 
moral principles to which the actor would (or should) conform. Finally, even 
if we could overlook the first two difficulties, it is not clear that Wertheimer's 
treatment of voluntariness is consistent with the rest of his theory.

First, are the moral principles that serve to delimit the scope of 
voluntariness those principles to which the agent does in fact adhere, or those 
to which he should adhere? If we take the former, subjective approach and 
argue that it is the agent's own principles that define the baseline for what it 
means to be "against one's will," this approach is clearly inconsistent with the 
law as it has developed. Certainly a judge, in determining whether a particular 

case constitutes an instance of duress, does not undertake an examination of 
the agent's moral principles—e.g., whether the agent is committed to prin-
ciples that could generate the range of existing alternatives from which he had 
to choose.

Moreover, this entire approach would relativize duress claims to each 
person claiming the defense. A person with an amoral set of principles which

61. I have argued in a previous piece that one dimension of autonomy is the vertical rela-
tionship between first and second order mental states, as Frankfurt argues. Another dimension of 
autonomy is the horizontal relationship between various second-order desires—i.e., that one's 
second order desires are consistent with one another so that they do not frustrate one another. John 
Lawrence Hill, Law and the Concept of the Core Self: Toward a Reconciliation of Naturalism and 
62. WERTHEIMER, supra note 1, at 9.
placed no constraints upon the just range of external options that should be open to an agent might be entirely precluded from maintaining a duress claim. On the other hand, for the person who holds a set of principles that limit the range of choice situations that are considered just, there would exist a correspondingly greater range of cases that the agent would characterize as coercive. And, of course, the great majority of persons have no developed set of principles whatsoever. Thus, the subjectivized approach not only fails to comport with existing law, it does so for good reason insofar as it would prove to be completely unworkable.

The objectivized approach is problematic for other reasons. First of all, what are the defining principles by which we are to judge voluntariness? Who sets the standard, and what is it? Moreover, even if we could agree upon some such objective standard, the objective rationalized account would have the implication that one's act could be involuntary without one knowing it, as when one acts in a situation that violates one's principles, but where one does not stop to consider this. Conversely, one could act voluntarily while believing his action to be involuntary, as where the actor is wrong about the application of these principles.

This leads to the second problem. The rationalized account of voluntariness is simply so divorced from the way in which we normally define voluntariness that it leads to anomalies such as those just mentioned. Most basically, there simply is no necessary link between one's being presented with a choice situation that violates one's moral principles and the voluntariness of one's act in responding to this choice situation. This cleavage between voluntariness and moral principle is particularly aggravated on objectivized accounts, where it is not clear in what sense one acts "against one's will" when one must act in contravention of principles that one should, but does not, hold. Most generally, the rationalized approach simply does not ring true as a matter of subjective experience. Every person knows the feeling that accompanies a coercive situation, and this experience has little or nothing to do with the moral principles to which one adheres. The point here is that any theory of duress which so thoroughly divorces the ascription of voluntariness from the phenomenal experience of coercion cannot be correct. Or, to put it differently, such a theory redefines coercion in a fashion that raises doubts about whether we are talking about the same concept. Wertheimer recognizes the problem in noting that the disadvantage of the rationalized account is that it operates by "detaching voluntariness from the will, from its psychological referent." 63

Finally, the third problem may be more aesthetic than logical. Nevertheless, it raises questions about the adequacy of Wertheimer's account as a whole. After devoting three-quarters of his book to an attack upon the traditional theory of duress and the corresponding notion that duress is a function of the voluntariness of an action, Wertheimer attempts to maintain that coercion is, after all, related to voluntariness (though, as we have seen, he rede-

63. WERTHEIMER, supra note 1, at 305.
The question that must be addressed is why does he see the need to attempt to relate the two concepts? The answer, I believe, is that we intuitively understand that coercion does have something to do with the etiology, rather than simply the immediate consequences, of the coerced act. It has to do with the fairness of holding persons responsible for acts performed in situations where their choices are limited. One is left with the unshakable impression that, as heroically as Wertheimer struggles to free coercion from the notion that it is essentially exculpatory in nature and function, he implicitly concedes the same in his treatment of voluntariness.

In conclusion, while there may be some convincing normative and linguistic reasons for distinguishing voluntary and free acts, such as the case of the heroin addict or the obsessive-compulsive, it is not because their (voluntary but unfree) behavior is inconsistent with some set of moral principles. It is because each is hopelessly conflicted, and because each cannot control the desire that leads to the resulting behavior. They act voluntarily in satisfying their respective desires to use a drug or to perform some repetitive action, but they have no control over the desire itself. They act voluntarily within the constraints of the situation, but they do not control or will the forces that shape that situation. In this respect, the victim of duress is analogously situated. The problem, however, is not a function of the moral principles to which each does (or should) adhere.

Part III

As we have already noted, the deepest irony underlying Wertheimer's theory of coercion is that it presents a moralized approach to coercion that provides no moral guidance as to when we should conclude that coercion exists. This is because, at bottom, the theory purchases descriptive accuracy at the price of a level of generality that precludes any prescriptive commitment. In essence, Wertheimer's theory approaches coercion from the outside, rather than from the inside. It looks more like the observations of a cultural anthropologist, or a social observer, than those of a philosopher performing applied ethical theory. His theory tells the judge that if she concludes that a proposal is immoral and that a given range of alternatives is impermissibly limited, she should find that the act was coerced. But he does not even presume to tell the judge how to answer these other questions—questions which themselves require the judge to have prejudged the coerciveness of the situation.

None of this is to say that moralized theories as a class are inherently flawed. The problem with Wertheimer's account is that it is prescriptively empty. The remedy for this fundamental difficulty is to construct a moralized account that specifies the prescriptive, rather than simply the descriptive, conditions for coercion claims.

What sort of a theory should this look like? First, it should reflect the essentially exculpatory nature of coercion. It should recognize, in other words, that we excuse the victim of duress because we understand that they are not responsible for the act, not because the act is justified in that it promote some good. At the same time, however, the exculpatory model should not be tied...
directly to questions of voluntariness, as with the traditional theory of coercion. Rather, the exculpatory nature of duress is a function of the futility of holding persons responsible for coerced acts. In sum, coerced acts are excused not because they are involuntary, but because they are not deterrable—because no appropriate legal sanction could, as a practical matter, have prevented the act.

The distinction between voluntary acts and deterrable acts highlights a central ambiguity with respect to the use of the term "voluntary." In the legal sense, as we have seen, it is sufficient for an act to be voluntary that the act resulted from a conscious mental state and a corresponding bodily movement by which the mental state (e.g., the intention) is carried out. Let us designate this first sense of the voluntary as "volitional" behavior. There is another, closely related sense of the term, however. This second sense of the term "voluntary" entails not only such volitional behavior, but requires additionally that the actor could have performed another course of action had she so chosen or, more generally, that other courses of action were open to the actor. Coerced acts arguably do not qualify as voluntary acts in this second sense insofar as other courses of action have been foreclosed by the coercive threat.

In discussing voluntary acts, we have employed the first sense of the word for three reasons. First, it comports with the technical legal sense of the term, as used in the criminal law, among other places. This avoids linguistic inconsistency brought about by using the term equivocally. Second, it is always open to argument that a person faced with a coercive threat could nevertheless have chosen to refuse to perform the act and incur the threatened harm. This would have the effect of collapsing the second sense of the term "voluntary" into the first insofar as the coerced actor can theoretically always resist the threat. Similarly, it raises a host of philosophical complications inherent in explicating the subjunctive concept that one "could have done otherwise"—issues which philosophers have found notoriously intractable.

Finally, where coercion is defined by the actor's inability to resist the threat, to say that an act will be considered voluntary only if the actor could have chosen otherwise would be to make the "voluntary" definitionally equivalent to the uncoerced. This would have the consequence of making the traditional theory of coercion tautologically true, since all coerced acts would be involuntary in this sense. For all these reasons, we have employed the first, legal sense of the term "voluntary" throughout this piece. Thus, an "involuntary" act, as we will use the term, is simply an act that does not issue from a conscious mental state through a corresponding bodily movement.

To say that an act is undeterrable is not to say that it is involuntary. If an act is deterrable, it is necessarily voluntary, but the converse is not true. Some voluntary acts are nevertheless undeterrable. Indeed, coerced acts are one species of the set of acts that are voluntary but undeterrable. (I leave open the possibility that some uncoerced acts might fall into this same class, such as acts motivated by some form of unyielding desire.) While conceptualizing duress as a function of deterrence casts the defense in explicitly utilitarian terms, there is a human side to this. It is precisely because we understand that coerced acts cannot be deterred that we empathize with the victim of coercion.
The utilitarian justification for duress reflects the very human, pre-analytic intuition that it is unfair to hold the victim responsible not simply because he did not engineer the coercive situation, but because the threat of punishment would not change his behavior. Thus, it is unfair to expect the victim of coercion to conform himself to the usual requirements of the law.

A utilitarian theory of coercion provides the determinacy necessary to guide judges’ decisions. The defense should be permitted in those cases where (1) but for the putatively coercive threat, the actor would not have performed the act, and (2) where the threat of certain punishment, appropriately relativized to fit the severity of the offense, would be ineffective in deterring the conduct. The second condition recognizes that deterrence is a relative phenomenon. This is important to make the defense available in cases where an inappropriately severe level of punishment would deter conduct for which the defense of duress should still be available.

For example, imagine a situation in which a street gang threatens to beat up a non-member unless he agrees to take part in vandalizing a school. If deterrence is measured from the standpoint of extremely severe punishment—say, the death penalty—it is likely that the non-member’s act is deterrable—i.e., it is likely that he would sooner take the beating than face execution for the act of vandalism. Thus, his act is theoretically deterrable and, as a result, would not be considered coerced if the second condition were not relativized. Holding the victim to this standard, however, would be unfair since his conduct was, by any ordinary definition, the product of coercion. Thus, the coercion claim must be made a function of a more appropriate level of severity of the punishment. As a rule of thumb, then, the second condition should be measured from the standpoint of the actual punishment threatened for a given offense in a just legal system.

The central organizing principle that undeterrable conduct should be excused explains and justifies a number of mediating rules which have developed at common law and more recently. It explains, for example, why those who recklessly place themselves in situations where they are likely to be coerced are denied the defense on the assumption that this prior conduct is deterrable. It also explains the modern trend in the law to permit the defense even for homicide, where the victim-defendant is threatened with the loss of his own life, or that of a loved one.

The foregoing remarks are obviously intended only as a general sketch for a utilitarian conception of coercion. Undoubtedly, a great deal more needs to be written to unpack fully the notion of deterrence central to the theory.

---

64. MODEL PENAL CODE § 2.09(2) (1962).
65. See supra note 24 and accompanying text for a discussion of the abrogation of the “no murder” rule under the Model Penal Code.
66. See Hill, supra note 6, at Part IV (discussing an account of the theory).