A Utilitarian Theory of Duress

John Lawrence Hill[*]

---

* Professor of Law, St. Thomas University School of Law. J.D., Georgetown University 1988; Ph.D., Georgetown University 1989. I would like to thank Lydia Chamberlain and Alex Nolte for their research assistance in writing this article.

INTRODUCTION

Duress is perhaps the most difficult defense to conceptualize and justify for a number of fundamental reasons. First, there is considerable disagreement concerning whether it is most accurately characterized as an excuse predicated upon the compelled nature of the act, or as a type of choice of evils justification analogous to the defense of necessity.[1] Second, if it is to be regarded as a type of excuse, it is arguably unlike all other exculpatory conditions in that it is the only excuse available to one who acts voluntarily, rationally, and with full information.[2] Moreover, its availability in both the civil and criminal realms creates problems with the consistency of scope and application in diverse legal areas.[3] Most generally, no other single excuse or justification has engendered as much controversy and debate with respect to its proper definitional characterization, its normative justification, and its legal application.[4]

The purpose of this Article is to defend a utilitarian account of duress against the two prevailing conceptions of duress: traditional or Aristotelian accounts, and so-called moralized or contextualized theories. In Part I, this Article begins by putting forth a 'theory of theories,' and elaborates a number of theoretical criteria from the philosophy of science which have been adapted and modified for evaluating the adequacy of competing jurisprudential theories. This will permit the comparison of diverse theoretical models with the object of determining which model best explains and justi-

---

1. 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 legally impermissible, the actor is not morally or legally responsible for its commission (excuse). *See* Kent Greenwalt, *Distinguishing Justification From Excuse*, 49 LAW & CONTEMP. PROBS. 89 (1986) (arguing that the distinction is far from clear). The Model Penal Code appears skeptical about whether the two concepts can be so neatly distinguished. MODEL PENAL CODE § 3 introduction. (1985). Perhaps the most important philosophical reason for this modern skepticism is that justification has been traditionally explicated as an act utilitarian defense, where the defendant can prove that the greater evil was avoided by his act, while excuse sounds in the paradigm of deontology. To the extent that excuse is also viewed in recent times in utilitarian terms, the two categories of defense are more difficult to distinguish. *See* H.L.A. HART, PUNISHMENT AND RESPONSIBILITY: ESSAYS IN THE PHILOSOPHY OF LAW 28-53 (1968) (discussing the utilitarian account of excuse). *See infra* notes 20-46 (discussing excuse and justification at greater length).

2. *See infra* notes 47-55 and accompanying text (examining the relationship between duress and voluntariness).

3. *See* ALAN WERTHEIMER, COERCION 19-53 (1987) (discussing coercion in contract law); *id.* at 144-69 (discussing duress in the criminal law). Wertheimer concludes that it is much more difficult to establish a claim of duress as a defense in a criminal matter, vis-à-vis a civil action. *Id.* at 153-54.

4. *See infra* Part IV.B. (defending the view that the utilitarian approach to duress is consistent with its exculpatory character).

fies the defense of duress. The remainder of the Article will employ these criteria in discussing the Aristotelian, moralized, and utilitarian theories of duress.

Part II surveys and critiques the traditional, Aristotelian theory of duress. The traditional theory, which greatly influenced common law judges and remains the pre-eminent conceptual explanation for duress in modern case law, holds that duress is to be understood as an excuse accorded in virtue of the involuntariness of the act. All theories that characterize coerced behavior as essentially unfree, or as the product of an "overborne will,"[5] partake of this model. This Article argues that coerced acts cannot be distinguished from other acts on the basis of voluntariness or freedom. Indeed, acts performed under conditions of duress are entirely voluntary, as the term "voluntary" is employed in other areas of the law. Moreover, attempts to conceptualize coerced acts as emanating from an overborne will presuppose a philosophical model of human motivation that has been seriously in question from the time of the late eighteenth century. In sum, coerced choices and acts cannot be distinguished from a variety of other situations involving constrained choice where no defense is available. For this reason, this Article argues that the traditional model must be rejected because it fails to justify the defense of duress.

In Part III, this Article considers so-called 'moralized' theories of duress, which have become increasingly popular among academic writers in recent years. In particular, it considers the partially and radically moralized theories of Joshua Dressler and Alan Wertheimer, respectively. Moralized accounts attempt to argue in various ways that coercion is not a function of the lack of freedom of the subject, or even the psychological pressure under which putative victims of coercion act.[6] In this respect, defenders of moralized accounts are also doubtful of the traditional model of coercion. They argue instead that notions of duress and coercion simply represent a kind of linguistic placeholder for more contextualized normative judgments that serve to exculpate or justify certain acts under certain situations. On this view, the concept of coercion is a kind of value judgment masquerading as a factual assessment of the moral-psychological state of the actor, i.e.,

---

5. In the context of duress as a defense in contract law, see generally *Kaplan v. Kaplan*, 182 N.E.2d 706 (1962); Austin Industrial Co. v. Loral Corp., 272 N.E.2d 533, 535 (1971); 13 SAMUEL WILLISTON & WALTER H.E. JAEGER, A TREATISE ON THE LAW OF CONTRACTS § 1605 (3d ed. 1970). For the classic statement, see *Silsbee v. Webster*, 50 N.E. 555 (1898) (stating that "duress going to motives consists in the threat of illegal acts").

In the criminal context, the Model Penal Code provides that duress excuses when "a person of reasonable firmness in his situation would have been unable to resist." MODEL PENAL CODE, § 2.09 (1962). The essence of the excuse is involuntariness. As one English case put it, a defendant should be acquitted "if the will of the accused [was] overborne by threats of death or serious bodily injury so [that] the commission of the offence was no longer the voluntary act of the accused." Regina v. Hudson 2 All E.R. 244, 246 (Crim. App. 1971).

6. *See* WERTHEIMER, *supra* note 3, at 7 (comparing empirical and moralized theories of coercion).

whether the actor was free or acted voluntarily. This Article argues that, in addition to a number of internal difficulties, moralized theories are hopelessly vague and possess little or no prescriptive value. They are high level abstractions which neither explain why coercive situations are viewed as exculpatory conditions nor guide judges, lawyers, or others in determining when a situation is coercive.

In Part IV, this Article explores the utilitarian model of duress. It begins by discussing the utilitarian theory of punishment and the corresponding rationale for exculpating certain acts. Next, it argues that duress is best understood as a kind of excuse, rather than a justification. This Article then endeavors to demonstrate that duress should be conceptualized not in terms of involuntariness, but as a function of the undeterrability of coerced acts. Here, this Article examines a number of issues in the law of duress that remain contested to which utilitarian theory can provide a solution. In general, it concludes that the utilitarian approach to the problem of duress will be more intellectually satisfying and more practically workable than the preceding two alternatives.

In Part V, this Article explores some objections to the utilitarian theory of duress. These will include objections that the utilitarian theory cannot generate conclusions with the degree of quantitative accuracy that it implies, that utilitarianism leaves out the retributive dimensions of criminal punishment and that a utilitarian approach may require punishment where punishment is not morally justified. Part V also evaluates the application of utilitarian theory to duress specifically. Finally, in Part VI, this Article will conclude with a brief discussion of the utilitarian approach to duress from the standpoint of the five criteria for theoretical adequacy outlined in Part I.

I. A THEORY OF THEORIES

Before addressing the issues noted above, an even more fundamental problem must be faced: what should be the criteria for choosing one theory over another? What renders one theory superior to others?

In the scientific realm, five criteria serve to evaluate the adequacy of theories. First, and most obviously, a theory must be relevant to the phenomena which it seeks to explain. Second, it must be testable, i.e., there must be some observable set of conditions under which the theory could be shown to be incorrect by failing to predict the outcome of some experiment. Third, it must be compatible with other previously accepted hypotheses. Fourth, it must possess predictive or explanatory value; it must be able to predict future observable phenomena and/or account for previously established facts. Finally, the theory must be simple; it must achieve the other four criteria with the least number of assumptions and in the least complicated way vis-à-vis other potential theories.[7]

7. *See* IRVING M. COPI, AN INTRODUCTION TO LOGIC 470-75 (6th ed. 1982) (discussing

Nonscientific theories, of course, differ in a number of fundamental ways from those that are scientific. Perhaps most importantly, the second criterion of a scientific theory, testability, cannot be interpreted literally in the context of a non-empirical theory. Jurisprudential theories of duress such as those to be examined here obviously cannot be disverified through experimentation by reference to some empirical state of affairs.[8] A theory of duress can, however, fail to comport with our normative intuitions regarding whether there ought to be a defense in particular cases. To the extent that a theory fails to conform to these judgments, i.e., where, for example, a particular theory has the consequence of casting the exculpatory net too broadly, thereby permitting a defense where one ought not obtain, or where the defense is too narrowly limited, we have some reason to question or reject the theory. This normative over or under-inclusiveness will not necessarily entail the rejection of the theory, of course. In some cases, rather than rejecting the theory, we might decide to change our pre-analytic normative intuitions, or to modify both the theory and the intuitions where necessary.[9] Nevertheless, a theory of duress that does not generally comport with our normative intuitions regarding the scope of its application should, under normal circumstances, be rejected.

Moreover, the third criterion of a scientific theory, compatibility with previously accepted hypotheses, must be re-interpreted in the jurisprudential context to require both descriptive accuracy, i.e., it must explain previous case law involving the issue of duress, and external consistency with other areas of law. For example, a theory that entailed that duress should be invocable as a criminal law defense whenever the actor subjectively feels a high degree of psychological pressure would be flatly inconsistent with the developed law, and should probably be rejected as descriptively inaccurate.[10] Similarly, a theory that does not serve to differentiate the function

---

these five criteria).

8. Jurisprudential theories differ from scientific theories with respect to the obvious fact that jurisprudential constructs are not empirical. Thus, jurisprudential concepts such as responsibility, excuse or duress are not amenable to empirical disverification. *See id.* at 471 (discussing testability). Indeed, disverifiability is usually viewed as a mark of the empirical, vis-à-vis the metaphysical. Put even more starkly, one cannot be said to have a theoretical hypothesis without there existing some possible opportunity for disverification, i.e., falsification by reference to some disconfirming observation. *See* CARL G. HEMPL, ASPECTS OF SCIENTIFIC EXPLANATION 39-46 (1965) (discussing absolute and relative conceptions of falsification and disverification from a logical positivist standpoint).

9. I have in mind Rawls' idea of reflective equilibrium here, by which theory and intuition may be mutually modified until we obtain a "fit" between the two. *See* JOHN RAWLS, A THEORY OF JUSTICE 48-50 (1971) (discussing the process by which we move from pre-analytic intutition to theory and back to intuition to make theory and our case-by-case intuitions correspond to one another).

10. *See, e.g.*, MODEL PENAL CODE § 2.09 (1985) ("It is an affirmative defense that the actor engaged in the conduct . . . which a person of reasonable firmness would have been unable to resist."). The civil law is more conflicted on this account. *See, e.g.*, S.P. Dunhom & Co. v. Kudra, 131 A.2d 306 (N.J. 1957) (rejecting the "objective" test). *Cf.* WERTHEIMER, *supra* note

and justifications of duress from those of other defenses,[11] or that is premised on principles of exculpation that are inconsistent with contemporary case law, will fail in virtue of the requirement of external consistency.

Finally, the fourth criterion of a scientific theory, predictive value, is not apposite in the context of non-empirical phenomena.[12] Rather, in a jurisprudential context, a parallel function to that of predictive ability is the capacity of a theory to guide judges in their determinations as to when duress does and does not obtain. Thus, rather than *predictive* value, a sound theory of duress should have *prescriptive* value. The prescriptive aspect of a theory provides the justificatory underpinning of the theory; as such, it has two related functions in the theory. First, it provides an organizing principle that serves to explain and systematize our pre-analytic normative judgments. It explains, in the case of duress, the reason or purpose for the defense; this also serves to limit the proper scope of the defense to cases where this purpose is satisfied. Second, this principle guides the judge's decision in particular cases. In so doing, the prescriptive aspect of a theory links our normative intuitions to the outcomes of future cases. In sum, the prescriptive component is the justificatory heart of the theory.

In addition to the criteria necessary to evaluating a scientific theory, another criterion is essential; this is the requirement that a theory be internally consistent. This condition, which is presupposed in scientific theory,[13] is explicitly elaborated here because of the expository nature of jurisprudential theory vis-à-vis scientific theory, and because of the greater likelihood of an implicit contradiction in jurisprudential analysis. Internal consistency requires that a theory should neither embody contradictory assumptions nor entail contradictory results or consequences.

3, at 33-34 (arguing that the civil law occasionally employs the objective test).

11. A "moralized" theory of duress, such as the one offered by Professor Wertheimer, may not be able to adequately distinguish between duress, on one hand, and unconscionability, enticement, and provocation, on the other. If duress is understood simply as a defense that recognizes the privilege of the victim to act in the face of an immoral proposal, WERTHEIMER, *supra* note 3, at 307, it appears that these other defenses cannot be distinguished, at least on this most general level of description.

12. Of course, jurisprudential constructs could be recast in empirical terms by reference to what a judge will do in a given situation. In this fashion, the judge's decision has a parallel function to that of an experiment in the scientific realm. Jurisprudential concepts would thus be operationalized by saying that a given situation was not an instance of duress if a judge says it was not. The problem with this, however, is that a judge's decision might be wrong. Moreover, this does not give the judge any assistance in the event that she is not certain whether duress obtains in a given instance. Finally, even if judges routinely disagree, it would not be incoherent for a third party to maintain that duress does in fact cover a situation where a judge has held to the contrary.

13. Internal consistency is presupposed by the criteria of testability and predictive power. A theory that embodies a contradiction would be capable of generating any outcome on the principle that anything follows logically from a contradiction. This, in turn, would undermine the testability and predictive power of any theory, because there would be no possible instance of falsification. In sum, internal consistency is necessary for disverifiability, itself a hallmark of a genuine scientific theory.

These five criteria for assessing the adequacy of a theory of duress—(1) normative testability, (2) descriptive accuracy, (3) internal consistency, (4) external consistency and (5) prescriptive power—will form the basis of discussion throughout this Article. Any mention of the relevance criterion has been omitted, the need for which is obvious and the function of which is fulfilled by the other criteria for the same reasons.[14] Also excluded is a discussion of simplicity.[15]

Even in the realm of science, competing theories might fare better or worse on different criteria, raising further questions concerning the way in which diverse theories are to be compared. This is particularly true in periods of scientific upheaval, as when a new theory possesses greater predictive power than a previously accepted theory but is inconsistent with a host of other previously well-established hypotheses.[16] This difficulty makes clear that, even in the realm of the empirical, qualitative judgments must be made in assessing the adequacy of competing theories.[17] This is even more the case with respect to non-empirical theories such as the various theories of duress to be evaluated here. For example, where case law has developed in some inconsistent manner, descriptive accuracy can only be purchased at the price of internal consistency.[18] Similarly, where previous case law fails to comport with our considered normative judgments, a theory cannot possibly entirely satisfy the elements of normative testability and descriptive accuracy.[19] For instance, any theory that refuses to permit the defense in a situation that appears clearly coercive or in violation of the victim's will may be problematic. These conflicts are virtually inevitable insofar as the case

---

14. For example, to meet the criterion of descriptive accuracy, the theory must be relevant to the topic described.

15. Simplicity is an application of the metaphysical principle of Occam's razor to the scientific domain. *See* BERTRAND RUSSELL, A HISTORY OF WESTERN PHILOSOPHY 472 (1966) (discussing Occam's razor).

16. *See* THOMAS KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS 77-91 (1970) (discussing the sociological aspects of paradigm change in science as an emerging theory challenges previously existing hypotheses).

17. This can be particularly problematic because theories tend to generate their own methodological assumptions as to evaluative criteria for a theory. Thus, there appear to be no objective, trans-theoretical assumptions for assessing competing theories. In quantum physics, for example, the methodological assumptions which underlie the concept of virtual particles could not be evaluated from the standpoint of the earlier Newtonian model. *See* TIMOTHY FERRIS, COMING OF AGE IN THE MILKY WAY 35-52 (1988) (providing a popular account of virtual particle theory).

18. In other words, any theory that could generate inconsistent conclusions must embody an internal contradiction. This confronts us with having to make the choice between adopting an inconsistent theory and having to ignore a portion of the existing law as incorrectly decided.

19. For example, the traditional duress defense in the criminal law never applied as a defense to murder. *See infra* note 188 and accompanying text (discussing the common law of duress and murder). If our normative judgments require that duress defenses apply even in cases of homicide, there would be a conflict between descriptive accuracy and normative testability.

law has developed without the guidance of an overarching theory that links normative intuitions, descriptive accuracy, and prescriptive power. Nevertheless, the fact that the five criteria cannot be applied in some rigorous, outcome determinative fashion does not gainsay their importance. In most cases, one theory will emerge as clearly superior to others, even if it is not superior in every respect to other competing theories.

As is demonstrated in the pages that follow, the utilitarian theory of duress is superior to the two prevailing theories in terms of all of the aforementioned criteria.

II. THE TRADITIONAL ARISTOTELIAN ACCOUNT OF DURESS

A. THE TWO-PRONGED THEORY OF EXCUSE

A threshold issue to the academic debate involving duress concerns whether duress is to be treated as a justification or as an excuse. As a general matter, the two are distinguished in that justification defenses are said to go to the act, while excuse defenses go to the condition of the actor.[20] More specifically, an act is justified when, on consideration of the facts leading to the act for which justification is sought, we wish to encourage such acts; an act is excused, on the other hand, when the act performed is otherwise morally (and legally) sanctionable, but where the individual is not responsible for the act.[21] Thus, self defense and necessity are classic justifications, the former because of the primal moral intuition that a person should be able to defend herself against attack without fear of criminal sanction, the latter because of the greater harm which would occur in the absence of the justified act.[22] On the other hand, acts precipitated by insanity, provocation, automatism and various diminished capacity conditions are excused because of the psychologically compromised state of the actor.[23]

While duress has traditionally been conceptualized as a type of excuse, more recent commentators appear fairly evenly divided on the question as to whether it should be treated as a type of exculpating condition analogous to physical compulsion, or as a justification similar to necessity.[24] Nor is the

---

20. *See* PAUL H. ROBINSON, CRIMINAL LAW 479 (1997) (arguing that justification permits a defense because the act is socially desirable, while excuses absolve the actor of responsibility).

21. *See* Paul H. Robinson, *Criminal Law Defenses: A Systematic Analysis*, 82 COLUM. L. REV. 199, 220-21 (1982) (contrasting justification and excuse).

22. *See* WERTHEIMER, *supra* note 3, at 144-69 (distinguishing necessity and duress by arguing that necessity is an agent-neutral justification while duress is an agent-relative justification); Joshua Dressler, *Exegesis of the Law of Duress: Justifying the Excuse and Searching for its Proper Limits*, 62 S. CAL. L. REV. 1331, 1347-49 (1989) (distinguishing necessity from duress); *id.* at 1356-67 (arguing that duress is better understood as an excuse, not a justification).

23. *See* JOEL FEINBERG, DOING AND DESERVING 1274-75 (1970) (discussing the Aristotelian bases for excuse); Michael Moore, *Causation and the Excuses*, 73 CAL. L. REV. 1091 (1985) (discussing these excuses).

24. The following commentators have argued that duress should be viewed as an excuse:

categorization debate a mere academic exercise. Deciding whether duress should be viewed as an excuse or a justification is important insofar as the underlying purpose and applicability of the defense will vary with excuse and justifications rationales, respectively.

For example, where duress is viewed to be a type of choice of evils justification analogous to the defense of necessity in civil and criminal law, it can only be used in situations where compliance with the coercive demand causes less harm than noncompliance.[25] Indeed, in defending the proposition that duress is a type of justification, a number of writers have pointed to the fact that the defense is not permitted at common law in cases of murder.[26] The defense is not permitted in these cases, the argument runs, because the harm with which the victim of coercion is threatened can never be greater than the consequence of compliance (i.e., the murder of an innocent third party by the coerced actor).[27] Where duress is treated as a type of excuse, on the other hand, the central question will focus upon the condition of the actor, and whether she should be held morally and legally responsible for the act, rather than on the overall beneficial consequences of the act. Part IV.B of this Article argues that the utilitarian paradigm treats duress as an excuse.

The modern Anglo-American paradigm of excuse is traceable to the writings of Aristotle.[28] In the *Ethica Nicomachea*, Aristotle puts forth a two-pronged theory of exculpation: an act is involuntary, according to the philosopher, if it "take[s] place under compulsion or owing to ignorance."[29] Adumbrating the modern philosophical divide between defects of cognition and defects of volition,[30] this view anticipates modern liberal notions which require that acts are morally attributable when and only when they are performed both rationally and freely.[31] With some modification, all modern

---

GEORGE FLETCHER, RETHINKING CRIMINAL LAW 830 (1978); Dressler, *supra* note 22, at 1350; Moore, *supra* note 23, at 1098. Others adopt the view that duress is a justification. *E.g.*, WAYNE R. LAFAVE & AUSTIN W. SCOTT, CRIMINAL LAW 433 (2d ed. 1986); WERTHEIMER, *supra* note 3, at 166; GLANVILLE WILLIAMS, CRIMINAL LAW 755 (1961).

25. *See infra* notes 159-64 and accompanying text (noting criticism of this limitation on duress).

26. *E.g.*, WERTHEIMER, *supra* note 3, at 166; Dressler, *supra* note 22, at 1352.

27. *See* LAFAVE & SCOTT, *supra* note 24, at 434 (2d ed. 1986) (discussing this problem and arguing that, for this reason, duress should be viewed as a justification).

28. ARISTOTLE, *Ethica Nicomochea*, *in* THE BASIC WORKS OF ARISTOTLE (Richard McKeown ed., 1941); *see* John Lawrence Hill, *Exploitation*, 79 CORNELL L. REV. 631, 655-59 (1994) (tracing the origins and development of the modern exculpatory paradigm from Aristotle).

29. ARISTOTLE, *supra* note 28, at 964.

30. FEINBERG, *supra* note 23, at 274-75.

31. This doctrine is most explicitly evident in the doctrine of informed consent, which requires that medical procedures shall be performed only after the patient has given her fully informed voluntary and rational consent. *See* PRESIDENT'S COMM'N FOR THE STUDY OF ETHICAL PROBLEMS IN MED. AND BIOMEDICAL AND BEHAVIORAL RESEARCH, SUMMING UP 17-22 (1983) (discussing the doctrine of informed consent in medicine); JOHN GRAY, LIBERALISM 56-60 (1995) (surveying the "idea of freedom" in both its positive and negative interpretations). *See generally* ALEXANDER BICKEL, THE MORALITY OF CONSENT (1975) (discussing the

forms of excuse, whether in criminal or contract law, are traceable to one or another of these two exculpatory prongs.[32] In sum, in order for an act to be excused, the proponent of exculpation must demonstrate either that the act resulted from some obstacle to effective reasoning or from some volitional failing, broadly construed.

The first prong of the exculpatory paradigm, excuses predicated upon defects of reason, takes one of two forms.[33] First, the exculpatory condition might be based upon a justifiable lack of knowledge on the part of the actor. This condition is reflected in defenses such as those predicated upon fraud or mistake in contract law, or mistake of fact in criminal law,[34] where the actor, through no fault of his own, does not possess certain information. The second class of exculpating conditions based upon the cognitive prong are what we would today classify as incapacitation defenses.[35] They are defenses based not upon lack of substantive information, but upon the inability of the actor to reason effectively. These include the modern insanity test pursuant to the M'Naughten rules, as well as various incapacity defenses in contract and criminal law. While these two sub-components of the cognitive prong for exculpation are not always separable,[36] they serve to distin-

---

political and moral implications of a moral system predicated upon consent).

32. Hill, *supra* note 28, at 655-59 (discussing the two exculpatory prongs in criminal, tort, and contract law).

33. Aristotle called these acts performed "owing to ignorance." ARISTOTLE, *supra* note 28, at 964. This included both what will be described in this article as genuine cognitive defects, and justified lack of knowledge. The second category of cognitive defenses are usually thought to limit freedom of the will, whereas the first does not insofar as effective use of the rational facilities is thought to be necessary to genuine freedom of the will. In defenses based on lack of information on the other hand, the actor is typically viewed to have acted freely, but without all of the necessary information. *See* FEINBERG, *supra* note 23, at 272-80 (discussing the various ways in which impediments to the cognitive process can effect exculpation).

34. These cognitive incapacity defenses include insanity involuntary intoxication in the criminal law and various defenses reflecting a party's incapacity to contract based on youthful age or mental disability.

35. Perhaps the clearest example is embodied in the M'Naughten insanity test, whereby a defendant is excused if he does not understand the nature or the wrongfulness of his action. M'Naughten's Case, 8 Eng. Rep. 718 (1843). Similarly, intoxication claims are generally admissible to negate specific intent, though not general intent, Avey v. State, 240 A.2d 107 (1968) (admitting an intoxication claim to negate assault with intent to kill) or for a diminished capacity defense, People v. Conley, 411 P.2d 911 (1966), or to negate malice, State v. Clark, 434 P.2d 636 (1967).

36. For example, incapacity due to a party's youth is characteristically viewed as a form of substantive incapacity to reason. Yet insofar as this type of incapacity is a function of lack of experience, it may be thought of as a lack of information type of incapacity. Conversely, M'Naughten-style insanity, which prohibits its victim from reasoning effectively, may function by limiting the victim's access to information, or by skewing his assessment of this information. Intoxication may prove the most difficult problem to categorize; on one hand, the intoxicated person may be viewed to have "forgotten," temporarily, information that normally forms the basis for our judgments. On the other hand, the intoxicated individual may simply exhibit poor judgment in failing to weigh such information. Other conditions, such as bipolar syndrome, which cause vacillating moods, may have a similar effect.

guish generally the two ways in which persons may fail to reason effectively.

The second prong of the traditional Anglo-American theory of excuse is the one with which this Article is more directly concerned. Generally, this prong entails that acts may also be excused when they are, in some sense, not within the volitional control of the agent. In this category fall those excuses predicated upon "no act" situations,[37] those that involve claims of genuine internal compulsion, such as the irresistible impulse test,[38] partially exculpating emotional conditions such as those brought about by provocation,[39] and a potpourri of (less successful) new excuses predicated upon physiological conditions, including the PMS defense,[40] the post-partum depression defense[41] and the XYY male syndrome.[42] All of these various (potentially) exculpating conditions are predicated upon the notion that the particular condition in some way poses an obstacle to the free and voluntary choice of the actor.

Notwithstanding recent academic criticism,[43] the traditional view of duress both in criminal and contract law was that coercion constitutes a species of volitional constraint pursuant to this second prong of the traditional exculpatory paradigm.[44] As such, duress is analogized to instances of

---

37. 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) (1962) 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. Similarly, in tort law, a volitional act is one requiring a mental element representing some motivational force, such as an intention or desire, and a corresponding bodily movement. W. PAGE KEETON ET AL., PROSSER AND KEETON ON THE LAW OF TORTS § 8, at 34-35 (5th ed. 1984) [hereinafter PROSSER & KEETON]. 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 nineteenth century, though it is evident in case law before then. *See* HART, *supra* note 1, at 97-98 (discussing Austin's view of the act requirement).

38. *See, e.g.*, Parsons v. State, 2 So. 854 (Ala. 1887) (serving, quite possibly, as the first case to specifically recognize this defense); FEINBERG, *supra* note 23, at 282-83 (raising doubts about the defense, or at least the modern application of the term "irresistible").

39. 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 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) (offering a limited defense of the EMED excuse).

40. *See* Robert Mark Carney & Brian D. Williams, *Criminal Law - Premenstrual Syndrome: A Criminal Defense*, 59 NOTRE DAME L. REV. 253 (1983) (discussing the medical and legal dimensions of the PMS defense).

41. *See* Marcia Barn, *Post-Partum Psychosis: A Legal Defense to Murder, or Both?*, 10 HAMLINE J. PUB. L. & POL'Y. 121 (1989) (reviewing cases and the medical evidence).

42. *See* LAFAVE & SCOTT, *supra* note 24, at 378.

43. A host of writers have rejected the "will" theory of duress. *E.g.*, P.S. ATIYAH, PROMISES, MORALS AND LAW 23 (1981); LAFAVE & SCOTT, *supra* note 24, at 433. *See infra* Part II.C (critically examining the various attempts to rehabilitate the will theory); WERTHEIMER, *supra* note 3, at 29.

44. The will theory is arguably reflected in the MODEL PENAL CODE § 2.09 () and in the

physical compulsion;[45] indeed, the use of the term "coercion" to represent cases of physical compulsion and instances of duress reflects the prevailing view that duress is an excuse which exculpates because the actor's will is "overborne" or because, most generally, the act was unfree insofar as the actor was forced to act against her will.[46]

The traditional view of duress requires viewing as essentially unfree acts performed under what are deemed to be coercive conditions. The problem with this view, as discussed in the next section, is that there appears to be no way of unpacking the concept of freedom or voluntariness which can consistently distinguish putatively coerced acts from acts performed in other presumably non-coercive situations.

B. THE PROBLEM OF VOLUNTARINESS

While a few commentators have maintained that coercion obviates the voluntariness of an act,[47] the weight of scholarly authority holds that an act performed under conditions of duress is nevertheless voluntary in the usual sense of the term.[48] Certainly coerced acts are voluntary in the conventional legal sense. Thus, in order to ground liability for an intentional tort, modern tort law requires that the defendant act volitionally: there must be both a mental state approximating some conscious intention or desire to bring about the act, and a corresponding bodily movement by which this mental state is manifested.[49] The actus reus requirement of the criminal law requires the same two components for an act to be considered intentional, reflecting the underlying mind/body dualism that underlies common law assumptions regarding human behavior.[50] If this is what is meant by acting

---

RESTATEMENT (SECOND) OF CONTRACTS § 492 (1981). Wertheimer argues, however, that the objective "person of ordinary firmness" standard enshrined in § 2.09 of the MPC effectively abandons the will theory. WERTHEIMER, *supra* note 3, at 34.

45. *See* Michael D. Bayles, *A Concept of Coercion, in* COERCION 17 (J. Roland Pennock & John W. Chapman eds., 1972) (comparing what the author calls "occurrent" and "dispositional" coercion, or physical compulsion and coercion, respectively).

46. Wertheimer, For one, argues that "[c]oerced choices are not unwilled, but they are, it may be said, against one's will." WERTHEIMER, *supra* note 3, at 302. As we shall see, however, it is difficult to see how any act that is voluntary is, nevertheless, against one's will. In one sense, where a person exhibits conflicting motives, any act which accords with one desire will frustrate the other. The victim of coercion may be viewed in this manner, e.g., she may have a desire not to do the coerced act, but she also has a desire, given the circumstances, to perform the act. Thus, it is difficult to see in what sense such an act is "against one's will." *See infra* Part II.C (discussing the seemingly insurmountable problem of answering this "problem of voluntariness").

47. *E.g.*, FELIX OPPENHEIMER, DIMENSIONS OF FREEDOM 15-17 (1961); BAYLES, *supra* note 45, at 17.

48. *See supra* note 37 (listing sources that review the conditions for volition in tort law and the actus reus requirement in criminal law).

49. *See* RESTATEMENT OF TORTS (SECOND) § 2 cmt. A (1965) (discussing volition as a component of the "act" requirement).

50. Ultimately, the modern act requirement embodied in the criminal law follows Des-

voluntarily or freely, then surely the person who acts under circumstances of coercion acts voluntarily, in both the conventional and legal senses of the term, insofar as she acts with the intention of bringing about the consequences demanded by the coercing party.

There is, however, a second sense of the term "voluntary" which must be discussed. The term is sometimes used with a subjunctive connotation to mean that a person could have acted in an alternative manner had she so chosen. In this sense, to say that an act is voluntary is to say more than that the act was volitional, as in the first sense. It is to say that, with other reasonable choices open to the actor, she nevertheless chose the course of action to which she ultimately committed herself. It is in this sense that an act is involuntary when it is "against one's will," and thus is deemed a coerced act thought to be involuntary. [51]

Proponents of the traditional theory of duress face a simple dilemma here. If the defender of the traditional theory argues that coerced acts are involuntary and uses the first (volitional) sense of the term, the argument is simply false. Coerced acts are, perhaps by definition, acts characterized by volitional responses to impossible choices. On the other hand, if "voluntary" is used in the second sense to require the existence of reasonable external alternative courses of action, there are other difficulties. Not only does this second sense of the word depart from the conventional legal sense in which the term is used in other contexts, but it requires highly nuanced normative judgments regarding the range of options open to an actor. This second sense of voluntary effectively necessitates abandoning the traditional theory in favor of some version of a moralized theory of duress. We shall return to this problem in Part II.A.

Aristotle was perhaps the first to struggle with problems of voluntariness and duress. His conclusion on the matter is notoriously ambivalent. He begins his discussion by considering both cases of what would today be classified as duress (e.g., where a tyrant threatens one with the death of her parents and children if she does not perform some base act) and cases of necessity (e.g., throwing goods overboard from a ship caught in a storm).[52] In the same paragraph, he states that such acts are "mixed, but are more like voluntary actions," because the "principle that moves the instrumental parts of the body in such action is in [the actor];" in another sense, however, such acts are "in the abstract perhaps involuntary, for no one would

---

carte's dualistic assumptions regarding the mind/body distinction. For an early elaboration of the act requirement, see John Austin, *The Province of Jurisprudence Defined*, Lectures XVIII-XIX. *See* RICHARD A. POSNER, THE PROBLEMS OF JURISPRUDENCE 161-96 (1990) (attacking this Cartesian ontology and defending a behavioristic view of human acts which dispenses with mental elements). For an emergent view of mind and its implications for law, see John Lawrence Hill, *Law and the Concept of the Core Self: Reconciling Naturalism and Humanism*, 80 MARQ. L. REV. 289, 344-55 (1997).

51. *See* WERTHEIMER, *supra* note 3, at 9 (discussing three senses of the term "voluntary").

52. ARISTOTLE, *supra* note 28, at 964.

choose any such act in itself."[53] In the end, he concludes that "pardon is [bestowed] . . . when one does what he ought not under pressure which overstrains human nature and which no one could withstand."[54]

If Aristotle is read to mean that a person should only be pardoned when a person *physically* cannot withstand to refrain from an act, then obviously the scope of that which is considered involuntary will be very limited. In this sense, it is unlikely that acts performed under duress meet the criteria for involuntariness. On the other hand, if some broader construction of the term "involuntary" is meant, what distinguishes the voluntary from the involuntary?[55] Moreover, if acts performed under duress are involuntary, then what about acts resulting from other precipitating conditions, such as provocation or a poor social background?

From Aristotle onward, defenders of the traditional view of duress have struggled to construct a moral-psychological theory that serves to distinguish duress from nonexculpatory precipitating conditions of human behavior. Simply put, theorists and philosophers have sought to provide a theoretical justification for the (morally and empirically) unlikely proposition that acts performed under duress are more analogous to true physical compulsion than they are to acts performed under circumstances involving limited, unattractive options for which no excuse lies.

C. FAILED EXPLANATIONS

There have been three types of moves made in attempts to demonstrate that acts performed under circumstances viewed to be coercive are involuntary or, in a more general sense, unfree. The first follows Aristotle in making the distinction between the voluntary and the involuntary contingent upon whether the precipitating cause of the action can be characterized as arising from an internal or external condition, respectively. The second move employs the offer/threat distinction in maintaining that acts flowing from threats are involuntary while those resulting from offers are voluntary. The third move involves a modern twist on an ancient theme characteristic of the tradition running from Plato through Kant to John Rawls: it attempts to make voluntariness a function of the underlying moral principles that guide (or should guide) one's behavior.

1. The Internal/External Dichotomy

Aristotle's dictum that acts should only be considered involuntary

53. *Id.* at 965.

54. *Id.*

55. Aristotle recognized this problem. In the *Ethica Nicomochcea*, he maintains that the involuntary is defined by the presence of an external motivating force. *Id.* He subsequently, however, distinguishes the voluntary from that which is chosen, arguing that the voluntary "extends more widely." *Id.* at 967. Thus, there is voluntary but unchosen behavior for Aristotle.

when "the cause is in the external circumstances and the agent contributes nothing"[56] provides a departure point for the first type of explication of the traditional theory of duress. As shown above, the clearest example of situations where the agent contributes nothing is seen in acts of genuine physical compulsion, where the defendant could not be considered to have acted at all.[57] In this paradigmatic case of compulsion, the moving force is genuinely external to the agent in a physical sense. Although contemporary linguistic usage links such "no act" situations to the qualitatively different situation involving duress by using the term "compelled" to describe both, in fact the victim of duress is "contributing something" to the situation in a way that the victim of physical compulsion is not. The victim of coercion does indeed act, in every legal sense of the word, albeit under circumstances of constrained choice.

One initial alternative to expand conceptually the realm of the external is to characterize as externally motivated all acts performed in response to an external stimulus, rather than requiring that the force itself be external. On this view, coercion provides a defense because the victim is confronted with some external state of affairs (e.g., a gun to her head) to which she is forced to respond. Obviously, however, this is much too broad. As Aristotle himself clearly saw, virtually every human action comes in response to some external state of affairs, including acts that are considered quite free.[58]

A more promising intermediate approach would expand the realm of the involuntary from the purely *physical* to include the *psychological* sense of the term.[59] This expansion of the metaphor of compulsion to include psychological factors employs a model of human motivation that depicts the coercive condition as one that overcomes the actor's usual capacity for freedom of choice. As such, the will of the coerced actor is viewed to be a passive mediating structure through which the will of the coercer is transmitted. This may at least partially explain why duress traditionally has been limited to situations where a person, rather than a natural force, constitutes

---

56. *Id.* at 965.

57. *E.g.*, where a third party, by sheer dint of greater physical strength, forces the defendant to perform some act, such as by squeezing the defendant's finger on a trigger.

58. According to Aristotle:

> But if some were to say that pleasant and noble objects have a compelling power, forcing us from without, all acts would be for him compulsory; for it is for these objects that all men do everything they do . . . . [I]t is absurd to make external circumstances responsible, and not oneself.

ARISTOTLE, *supra* note 28, at 965-66.

See e.g., MODEL PENAL CODE § 2.09 (1985) (permitting duress in caes where a person of "reasonable fitness" would be "unable to resist"). This language makes duress appear to excuse literally because of a kind of compulsion operating at the psychological level.

59. This move is central to the "overborne will" theory prevalent in legal analysis. *See supra* note 5 and accompanying text.

the coercive stimulus.[60]

This expanded sense of compulsion, however, runs up against notoriously intractable difficulties. Indeed, the same fundamental problem considered in the previous section—the problem of distinguishing duress from other conditions for which no excuse is available[61]—is resurrected at another level. Where coerced acts are characterized as those in which the will of the coerced actor is rendered passive, it becomes difficult to distinguish cases in which a person acts because of threat of harm from those where she acts because of compelling desire. Moreover, cases in which acts are motivated by compelling desire may be difficult to distinguish. And how are we to distinguish coerced acts from those that result from external conditions that serve to shape the will of the actor, such as those embodied in the process of socialization?[62]

One way to preserve the traditional model of psychological compulsion is to maintain that acts performed under duress are externally motivated in the sense that the usual process of reasoning is short-circuited. For example, it may be argued that the person who acts in response to a life-threatening stimulus, such as having a gun at her head, may be said to act from some instinctual sense of self-preservation that precludes the normal deliberative and rational aspects of decision-making. In this sense, the act appears more analogous to a type of reflex than to a carefully chosen course of action.

There are, however, numerous problems with this view. First, duress is available in numerous situations where the self-preservation of the coerced actor is not threatened, as when the lives of third parties unknown to the victim are at stake, or where what is threatened is a less serious physical assault upon the person of the coerced actor. Indeed, the argument from self-preservation appears to have little to do with application of the duress doctrine in areas outside criminal law; in contract law, a threat involving some lesser interest may be sufficient to claim duress.[63] Secondly, and per-

60. As Bayles has argued, *coercion* requires "an interpersonal relation involving a complex intention on the part of a coercer." BAYLES, *supra* note 45, at 17.

61. In other words, the overborne will theory assumes that the will exists ab initio with a set of values and preferences intact. More recent views of human personality, however, point to the fact that personality is formed, that it develops as a response to external social influences. If this view is even partially accurate, then external influence can have an even more direct role on choices—i.e., not by overwhelming the existing will, but by forming the nascent will. *See generally* Richard Delgado, *"Rotten Social Background": Should the Criminal Law Recognize a Defense of Severe Environmental Deprivation?*, 3 L. & INEQUALITY J. 9 (1985) (providing a classic statement of the position that social influence affects the development of the will and that this should be legally recognized through a separate line of defenses).

62. How, for example, do we distinguish coercion, which operates by threat, from inducement, which operates by its affects upon powerful desires. Indeed, this is why seduction sometimes appears "coercive." More generally, if the will is a product of social forces which "construct" what we value and desire, as some maintain, then the will of the actor is always simply a medium for transforming social forces into human action.

63. This is more likely in the contracts context. *See* JOHN C. CALAMARI AND JOSEPH M.

haps more fundamentally, decisions made under circumstances of duress are altogether rational. Even where the life of the coerced actor is at stake, the victim acts in an entirely rational way, i.e., in handing over his wallet rather than facing death. Indeed, the very fact that coercion is characteristically viewed as causing a *volitional*, rather than a *rational* impairment, underscores our recognition of the fact that coercion is typically not viewed as overcoming the rational capacity of the actor.[64]

Yet another way of drawing the internal/external dichotomy along psychological lines that preserves some sense of the concept of psychological compulsion is suggested by Professor Michael Moore. We should only hold persons responsible, he argues, when they act from internalized values, goals and beliefs.[65] We distinguish the victim of duress from the product of a bad environment in holding only the latter responsible, he argues, because the latter acts from beliefs, desires and values that are truly her own in a genuine psychological sense. What distinguishes such "internalized" sources of motivation is that they have been made a part of the actor's character; they have been incorporated or "integrated" into the personality of the person over time, as the result of conscious choice. On the other hand, the victim of duress acts from "external" sources of motivation in the sense that she would not choose to perform such an act on her own; she does not desire to do that which she is compelled to do.

While there undoubtedly is a genuine subjective, psychological difference between the actor who is forced to act under coercive conditions and the actor who acts in a certain manner "because she really wants to," it is not clear that there exists any difference from the standpoint of the volitional capacity of the actor. Indeed, while I have defended the idea of internalization as a way to preserve the traditional model in a previous article,[66] I now find it indefensible. The basic problem is that there is no way to distinguish the motives of the victim of duress from those of non-coerced actors with respect to the internality of these motives. Though acting under conditions where external options may be greatly limited, the victim of duress desires and chooses her act as does the uncoerced individual. Indeed, particularly when acting from motives of self - preservation or the preservation of loved ones, the coerced individual acts from values that lie at the heart of human personality, values at least as "internal," in any psychological sense, as acts motivated by other desires.

PERILLO, THE LAW OF CONTRACTS §§ 9-5 and 9-6 (3d ed. 1987) (dealing with duress of goods and economic duress). *Cf.* LAFAVE & SCOTT, *supra* note 27, § 5.3(b), at 436 (threat of property damage generally not enough in criminal cases). Even here, however, the broader test of the Model Penal Code might permit a defense where a person of "ordinary firmness" would act under threat of property damage. MODEL PENAL CODE § 2.09 (1962).

64. *Accord* JOSHUA DRESSLER, UNDERSTANDING CRIMINAL LAW 277-78 (2d ed. 1995).

65. Moore, *supra* note 23, at 1136 (suggesting that the victim of brainwashing should be excused only if she did not have adequate time to integrate the experience into her existing belief structure).

66. Hill, *supra* note 50, at 377-81.

At a deeper theoretical level, the attempt to distinguish the victim of duress from the victim of a "rotten social background," or the person who acts from other, presumably "free" motives, appears to employ a questionable model of human behavior.[67] The problem is that there is apparently no principled way to distinguish the internal from the external at all. For example, physical conditions that disrupt normal behavior patterns, such as hormonal imbalances, genetic conditions or nervous disorders, may be constructed alternately as internal to the person insofar as they arise physically within the person, and as external conditions insofar as they are abnormal or irregular states.[68] To the extent that a person adopts and "internalizes" various desires, motives, values, and beliefs in response to her perception of the available range of alternatives, her internalized preferences and values may be as much a function of external options as is the desire to avoid the coercive stimulus on the part of the victim of duress.

The argument here is not that there is no psychologically or morally relevant distinction between the victim of duress and the person who acts from other motives. Indeed, as I argue in Part IV, there are important moral and legal reasons for distinguishing these cases. Rather, from the standpoint of the motives or the voluntariness of the actor, there appears to be no principled way to distinguish cases of duress from other, non-exculpatory precipitating conditions along psychological lines. If the traditional view of duress is to be salvaged at all, the voluntary and the involuntary must be distinguished in some other way than by reference to a continuum of internal and external conditions.

2. Threats and Offers

The distinction between offers and threats underscores the contemporary debate regarding coercion, at least within the liberal tradition. According to the classical liberal view, coercion is distinguished from non-coercive conditions by reference to whether the putative victim is sub-

---

67. The point is not simply that all human actions are causally determined, as behaviorists and other determinists maintain. *See, e.g.*, B.F. SKINNER, BEYOND FREEDOM AND DIGNITY 24-46 (1971) (arguing that human freedom, as in freedom of the will, is illusory); John Lawrence Hill, *Mill, Freud and Skinner: The Concept of the Self and the Moral Psychology of Liberty*, 26 SETON HALL L. REV. 92, 154-66 (1995) (discussing the implications of behaviorism for both our conceptions of personal responsibility and overall political liberty). Some determinists, those of a soft determinist orientation, argue that qualitative moral distinctions can be drawn between different types of precipitating causes of human behavior. Soft determinists include THOMAS HOBBES, LEVIATHAN 127-28 (Penguin ed. 1968) (1651); DAVID HUME, AN INQUIRY CONCERNING HUMAN UNDERSTANDING 90-111 (Charles Hendel ed. 1955) (1758); JOHN STUART MILL, A SYSTEM OF LOGIC 413 (1843); A.J. AYER, LANGUAGE, TRUTH AND LOGIC (1976).

68. *See* Hill, *supra* note 50, at 44, 299-303 (discussing this problem as it relates to physiological conditions, physiologically induced emotional states, environmentally caused physical states, and conditions the status of which is unclear in that they can be viewed either as physical or mental states).

ject to an offer or a threat. According to the liberal view, offers can never be coercive; only threats coerce.[69] In what way, however, are threats freedom-limiting in a manner that offers are not?

Threats may be viewed to be freedom-limiting in one of two ways: by their effect on the will of the actor, or by their impact on the range of external contingencies or choice options open to the actor. One version of the theory holds that offers and threats are distinguished in that threats appeal to sources of motivation over which the agent has less control, e.g., fear, self-preservation, while offers evoke only desires.[70] In other words, threats undermine the voluntariness of the act in a way that an offer never could because threats cannot be refused in the way that offers can.

This theory raises issues similar to those discussed in examining the concept of internalization. Most generally, there does not appear to be a psychological distinction, from the standpoint of voluntariness, between fear, desire and other motivating conditions. Even behavior motivated by fear is typically characterized as free, as when the agent escapes some dangerous condition. Thus, being motivated by fear or the instinct for self-preservation is not sufficient for a claim of involuntariness. Indeed, it is not clear how behavior motivated by desire, fear or any other conscious motivation could ever be involuntary in the sense that it is not volitional.

---

69. In this respect, the liberal tradition can be traced to Hobbes, who argued in favor of a negative conception of freedom, i.e., that freedom consists in the absence of external constraint. THOMAS HOBBES, LEVIATHAN, CH. 21 (C.B. MacPherson, ed. 1981) (1651). *See* ISAIAH BERLIN, *Two Concepts of Liberty, in* FOUR ESSAYS ON LIBERTY (1969) (discussing the positive and negative distinction and defending the negative conception of freedom). The negative conception of liberty is linked to the view that only threats can coerce via the assumption that offers only make additional choices possible, and thus can never constitute an external obstacle or constraint. *See generally* ROBERT NOZICK, *Coercion, in* PHILOSOPHY, SCIENCE AND METHOD (Ernest Nagel ed. 1972). The classical liberal assumption, of course, is that internal motivating forces, such as strong desires or emotions, cannot be viewed as creating an external obstacle because they are, by definition, internal. Interestingly, however, we might ask what makes threats any more freedom-limiting than offers. If, for example, what makes a threat compelling is that it motivates by fear, this is as much an internal motivating force as is desire, and so it is not freedom-limiting in the negative sense. On the other hand, if the liberal argues that threats limit freedom by reducing external options, then he is faced with two problems. First, many ordinary events limit our external options in an everyday sense. Yet the liberal does not view all such events as freedom-limiting. If the liberal introduces some qualitative distinction—i.e., that external options must be reduced in some extreme fashion—then this raises questions about how to draw the line and why only extreme cases should be freedom-limiting. Second, and perhaps more fundamentally, it is not immediately clear how reducing external options renders one unfree in pursuing the remaining options, at least on a liberal account of negative freedom as absence of constraint. *See* HAROLD D. LASWELL & ABRAHAM KAPLAN, POWER AND SOCIETY 97 (1950) (arguing that offers coerce); VIRGINIA HELD, *Coercion and Coercive Offers, in* COERCION (J. Roland Pennock & John W. Chapman 1972) (same).

70. Threats and desires may both motivate in an equally effective manner, but desires are more "internal" in the sense that they reflect the actor's own values and dispositions. The problem with this is that many acts can be characterized as resulting from either a desire or a fear—e.g., the fear of death or the desire to preserve one's life.

On the other hand, if by "involuntary" we mean the second subjunctive sense of the term, then serious threats in constrained circumstances may indeed be involuntary. This move, however, raises issues concerning the limited or constrained external circumstances, not the internal capacity of the actor.

Turning now to these issues, the second way of distinguishing (presumably freedom-limiting) threats from (freedom-enhancing) offers is by reference to the way in which each affects the range of external options of the actor. Most basically, a proposal is a threat when it reduces the available options open to the actor, while an offer, by definition, can never reduce the actor's options, and will usually increase them.

Putting aside for the moment the claim made by some that offers can sometimes be coercive,[71] even those within the classical liberal tradition who otherwise wish to defend the threat/offer dichotomy and the related Hobbesian conception that coercion requires external constraint, have recognized some of the difficulties with the position.[72] First, by a simple twist on normal linguistic usage, offers can be converted into threats and threats into offers. For example, rather than characterizing a job offer as such, it can be viewed as a threat on the part of the employer to the effect, "Come work for me or I won't give you the money."[73] Similarly, every employee coerces his employer by threatening not to work for the employer unless the employee is paid.[74] Conversely, all threats can be reduced to offers—e.g., the robber offers the pedestrian his life in return for his money.[75]

None of these examples, of course, refute the liberal dichotomy between threats and offers. All the defender of the distinction needs to do is to make clear that something will only be defined as a threat if it in fact reduces the range of available alternatives in comparison with the actor's ex ante situation. For example, the robber coerces her victim by effectively reducing his options vis-à-vis the pre-robbery encounter. Prior to the hold-up, the pedestrian can either retain his money or dispose of it in any number of different ways. His range of options is reduced to one, however, as a result of the encounter with the robber. In sum, what makes the robbery coercive is not the motivational source to which the robber appeals; rather, it is the reduction in the victim's available range of choices which is freedom-limiting.

---

71. HELD, *supra* note 69.

72. NOZICK, *supra* note 69, at 452. First, even in some non-threat situation, it might be legitimate to say the victim had no choice. *Id.* at 447. Second, there are some offers which are closely tied to threats, as where the coercer offers not to turn over to the police information that the coerced has committed a crime. *Id.* at 452. Similarly, someone might offer not to assist a coercer in bringing about a threatened consequence. *Id.* Finally, sometimes offers are so attractive that a person cannot reasonably be expected not to go along with it. *Id.* at 460.

73. *Id.* at 447; Dressler, *supra* note 22, at 1337 (noting that the offer/threat dichotomy relates to the fear/desire distinction, e.g., the robber tempts the victim with her life).

74. NOZICK, *supra* note 69, at 447.

75. Dressler, *supra* note 22, at 1337.

The problem with this view is that we do not, in other circumstances, hold that a reduction in external contingencies results in involuntary behavior. Having fewer options from which to choose—even where the range of options is radically reduced—does not render the choice itself involuntary, even if it might otherwise be unfair to punish the actor for the choice. Indeed, this view appears to conflate internal and external conceptions of freedom where, for example, it confuses freedom of the will with the existence of available external options open to the actor. The traditional paradigm excuses because of the effect upon the actor's moral psychological capacity to choose, not because she chooses from a limited range of alternatives. In sum, while the availability of limited options is relevant to duress, as Part IV argues, it is not because the actor has acted involuntarily.

Robert Nozick defends a version of the traditional model by employing a different twist on the offer/threat dichotomy.[76] In his essay on coercion, he maintains that "[w]hen a person does something because of threats, the will of another is operating or predominant whereas when he does something because of offers this is not so."[77] Thus, the person who acts from threats does not act voluntarily.[78] His response makes internal freedom or voluntariness a function of external freedom, or the availability of choices. He explains that:

> [T]he crucial difference between acting because of an offer and acting because of a threat vis-à-vis whose choice it is, etc., is that in one case (the offer case) the Rational Man is normally willing to move or be moved from the presituation to the situation itself, whereas in the other case (the threat case) he is not. Put baldly and too simply, the Rational Man would normally (be willing to) choose to make the choice among the alternatives facing him in the offer situation, whereas normally he would not (be willing to) choose to make the choice among the alternatives facing him in the threat situation.[79]

In sum, what makes threats freedom-limiting is our desire not to have them made at all.

Nozick's solution will not do here. Not only is his underlying premise false, but even if it were true, it would be irrelevant to the question of voluntariness. First, it is simply not true that offers are characterized by the offeree's willingness to have the offer made, as anyone who has had to contend with the unwanted solicitation of sales personnel can attest. If the mere unwillingness to have a proposal, whether an offer or a threat, is suffi-

---

76. NOZICK, *supra* note 69.

77. *Id.* at 459.

78. "[A] person who does something because of threats does not perform a fully voluntary action, whereas this is normally not the case with someone who does something because of offers." *Id.*

79. *Id.*

cient to preclude the ascription of voluntariness on the part of the actor, then certainly the scope of the involuntary will be dramatically enlarged.

More importantly, it is not at all clear that the unwillingness of the agent to have the threat made has anything to do with the voluntariness of her response to the threat. To put it somewhat differently, one's psychological disposition to a proposal of any kind appears to have little if any relevance to the moral quality—the voluntariness, freedom or culpability—of one's ultimate response to the proposal. While Nozick does not explain why the *ex ante* willingness or unwillingness of the actor to have the proposal is relevant to voluntariness, two possibilities suggest themselves. First, Nozick states in the above quoted passage that the Rational Man would normally "not (be willing to) choose to make the choice among the alternatives facing him in the threat situation."[80] Obviously, however, we do not normally conclude that a given act is involuntary simply because the actor is forced to choose among a range of alternatives with which she would not choose to be confronted. People are faced with such choices every day and are not excused for the consequences of their decisions.

Though Nozick does not elaborate on it, there is a second possible interpretation to the link between one's not choosing to confront a certain range of choices and one's voluntariness. This interpretation, however, represents an entirely different approach to solving the problem of voluntariness, which will be discussed next.

3. Voluntariness and Moral Principles

The rationalist tradition in Western philosophy offers another possible way to salvage the traditional model of duress. Recent thinkers including Harry Frankfurt,[81] Gary Watson,[82] and Charles Taylor[83] have sought to distinguish the merely voluntary from the truly free or autonomous by arguing that reason plays a special role in the moral psychology of decision-making. These thinkers reject the empiricist notion of freedom as unconstrained choice which has influenced modern thought from the time of Hobbes.[84] On this rationalist view, autonomy requires more than simply

---

80 *Id.*

81. Harry Frankfurt, *Freedom of the Will and the Concept of a Person*, 68 J. PHIL. 5 (1971).

82. Gary Watson, *Free Agency*, 72 J. PHIL 205 (1975).

83. CHARLES TAYLOR, WHAT'S WRONG WITH NEGATIVE FREEDOM, PHILOSOPHICAL PAPERS (1992).

84. The empiricist view of freedom is that which is associated with negative liberty or freedom as non-constraint. Thomas Hobbes was the original defender of this view. THOMAS HOBBES, LEVIATHAN 189 (C.B. MacPherson ed. 1968) (1651). A person is free if he is not prevented from acting in accordance with his desire, choice or preference. It tends to equate the free with the voluntary, as this term is used in conventional legal doctrine. This view is most closely associated with the empiricist tradition because it approaches freedom from an external of behavioristic standpoint. Most fundamentally the empiricist view rejects the notion that there is a metaphysical entity known as the will which mediates between competing desires, picking the best or most rational course of action. On crude empiricist accounts, there is

acting from desires. It requires that these desires themselves conform to considered moral judgments, principles that emanate from self-reflection and that serve to guide and sometimes restrict desire-driven behavior. This reflection vertically grounds our actions in principles that embody our deepest moral convictions, but it also horizontally rationalizes our desires, serving to make them consistent with each other. On this view, actions which are merely voluntary, which are performed in accordance with "first order" mental states[85], do not necessarily reflect stable preferences, or may result from heteronomous motivational factors. Thus, the rationalist argues that acts performed as the result of intoxication, seduction, addiction, or provocation, among others, may meet the empiricist definition of freedom but are not autonomous in any meaningful sense.

Wertheimer contends that duress should be conceptualized in a similar fashion. As he puts it, "the 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."[86] (emphasis in original). More specifically, a person 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)[87] affirm. On this rationalized view, co-

---

nothing but desire, and all acts are entirely desire-driven. Decision-making is nothing but the process by which a more powerful desire wins out over a less powerful desire. See, e.g., David Hume, for a slightly more sophisticated version of this model.

In contrast, rationalist accounts argue that reason or the rational will mediates between competing impulses. Acts can fail to be free when this reason-based process is short-circuited, even where the act is voluntary. One of the most interesting differences between these two accounts is that the empiricist tends to reject the idea of weakness of will. We always act, on the empiricist account, in accordance with our strongest desire. Thus, coercion, seduction, provocation and other potentially exculpatory conditions are treated differently by rationalists and empiricists.

85. Professor Harry Frankfurt adopts the distinction between first and second-order mental states. First-order mental states are thoughts, desires, beliefs and other intentional mental states that have as their objects other things—e.g. a desire for fame, a belief in one's abilities, etc. Second-order mental states are mental states about other mental states. According to Frankfurt, autonomy requires that our first-order states comport with out second-order states. For example, if one has a first-order desire for tobacco and a second-order desire that he does not desire tobacco, the subject's personality is in conflict and his act is not truly free. Frankfurt, *supra* note 81, at 210.

86. WERTHEIMER, *supra* note 3, at 301. 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.

87. Wertheimer waffles here on whether the defining principles are those the victim *does* or *should* affirm. At one point Wertheimer argues that it is the principles to which the victim is committed, and not the way he understands them. *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,

erced acts are not unwilled, they violate one's will in the sense that they do not conform to one's system of considered moral judgments.[88] An act is not coerced simply because the actor does not prefer to act in a certain way. As Wertheimer puts it, "reluctance and voluntariness can well go hand in hand."[89] An agent is coerced, however, when this preference not to act is based on her deepest moral convictions.

A similar Kantian twist can be applied to Nozick's claim that an act is involuntary when the Rational Man would not choose to be faced with the given range of alternatives.[90] This can be interpreted as holding that acts are involuntary when the person who would act autonomously, in the strict Kantian sense, must choose from a range of alternatives that *are against her rational will,* i.e., that violate the principles of morality and justice to which rationality commits her. Whether Nozick in fact meant anything of the sort is not clear, but this is perhaps the most straightforward interpretation of his theory.

The rationalized account of duress is hopelessly inadequate either to explain the phenomenal or subjectively experienced aspect of coercion, or to justify the defense of duress from a normative standpoint. The first problem with Wertheimer's version of the rationalized account is that he fails to make clear whether the principles to which we must appeal in determining whether a particular act is coerced are principles to which the actor does in fact adhere, or whether they are objectively given and, thus, are principles to which she *should* adhere. There are problems, moreover, with both such versions of the rationalized account. To put it most generally, subjective rationalized accounts appear to relativize coercion claims to the varying set of principles held by each particular actor, while objective accounts relinquish the connection to the will or to voluntariness necessary to any version of the traditional theory of duress.

On subjective accounts, a person is deemed to be coerced if she is forced to act in situations which violate her own set of principles (i.e., where she is forced to respond to a proposal and to choose from among alternatives that do not conform to what she subjectively views to be morally just or permissible). The anomalies generated by this should be evident. A person with very exacting principles will be able to claim coercion frequently while the person who holds relatively more loose principles will often be precluded from making such claims. More bizarre still, the person without any such considered principles—and this description may well fit the great majority of all persons—will have no basis to make any coercion claims whatsoever insofar as she does not hold any such considered moral judgments by which we are to gauge coercion claims.

---

as I argue in the next few paragraphs.

88. *See id.* at 301 (discussing the choice prong on this analysis); *id.* at 302 (discussing the proposal prong).

89. *Id.* at 302.

90. *See* NOZICK, *supra* note 69, at 459.

Objective rationalized accounts, on the other hand, sever the link between voluntariness and coercion. It is not clear in what sense a person may be said to act against her own will when she acts in situations where she is required to respond to a proposal and to choose from among some set of choices that fall below the minimum baseline established by a set of externally established principles. Indeed, on this view, a person may be coerced without even knowing it, as when she is not aware of the objective principles by which coercion is to be defined.[91] At the end of his book, Wertheimer acknowledges the problem of detaching the moral from the psychological aspect of coercion without addressing it.[92]

The traditional theory of duress fails to satisfy a number of the five criteria used for evaluating the adequacy of any theory. These failings flow largely from the difficulties embodied in the problem of voluntariness. First, the internal consistency of the traditional view is seriously in question. To the extent that the theory defines voluntariness as volitional behavior, or the ability to act, it cannot consistently hold that acts committed under duress are not voluntary. Nor will it do to employ a terminological subterfuge, for example, by distinguishing the involuntary from the nonvolitional.[93] The point is that the Aristotelian paradigm of excuse, from which the traditional theory flows, permits exculpation where there is something amiss with the agent's capacity to engage in free choice, not where the agent is faced with an unattractive range of external options from which to choose.

For similar reasons, the traditional view lacks external consistency insofar as the definition of voluntariness employed in the context of duress conflicts with the use of the term with respect to other exculpatory issues. Either voluntariness requires simply volitional behavior simpliciter or it requires something more. If it requires merely volitional behavior, then cases of duress are voluntary. If something more is required, then, not only does this appear to relinquish the traditional view of duress insofar as we must appeal to some factor other than voluntariness, but our new definition is quite distinct from the definition employed in other legal contexts. In this fashion, the criterion of external consistency is violated.

Nor can the traditional theory of duress provide prescriptive accuracy. The voluntariness criterion does not justify the defense as it has evolved in Anglo-American jurisprudence because acts committed under duress are voluntary in the usual legal sense of the word. Moreover, for this same reason, it cannot by itself guide judges in future decisions.

The traditional theory of duress must be relinquished because, when

---

91. This is also likely to be true in the situation where one does not understand the implications of the principles he or she endorses, believing that they have been violated when they in fact have not. *See* WERTHEIMER, *supra* note 3, at 302-03.

92. *Id.* at 305.

93. As Wertheimer points out, if there are volitional but not voluntary acts, one must demonstrate how some conscious motivation can nevertheless be involuntary. *Id.* at 304.

we define voluntariness as a manifestation of the inner moral-psychological capacity to do what one chooses under the circumstances, acts committed under duress are voluntary. On the other hand, when voluntariness is defined by reference to the availability of attractive external options, there is no clear way to draw the line between acts resulting from duress and those resulting from social conditioning, economic exigency, or related precipitating factors. Indeed, there is no systematic principle by which to distinguish the coerced from those who are faced with a range of options they find unattractive.

Still, the traditional view appears to approximate our pre-analytic intuitions regarding whether it is just to hold persons responsible for acts committed under duress. As Joshua Dressler, and H.L.A. Hart before him, argued, we excuse cases of duress because the actor has no "fair opportunity to choose."[94] This move explicitly incorporates into the duress analysis a normative criterion which represents a second approach to the problem of duress.

III. MORALIZED THEORIES OF DURESS

The traditional theory of duress is attractive in large part because it conceptualizes the distinction between voluntariness and coercion as an empirical one rather than a moral judgment. To say that these questions are empirical is not to maintain that they are easily verified.[95] Indeed, on particularly rigorous accounts of what it means to be empirical, such as those of the logical positivist, questions about mental states might not be empirical at all in the sense that they are not inter-subjectively observable.[96] Nevertheless, questions of voluntariness are arguably empirical in nature insofar as they reflect psychological facts concerning the etiology of certain behavior. This is so even if these psychological states can only be described and verified on a first-person basis by the actor herself. Moreover, before too long, there may come a day when science is able to verify objectively such states by reference to observable brain processes.

For a variety of reasons, however, factual accounts of duress have proven untenable. As shown in Part II of this Article, the most likely factual

94. Dressler, *supra* note 22, at 1365. *See also* HART, *supra* note 1, at 45 (arguing that mistake, coercion, ignorance, and undue influence, among others, are excused because the individual moved by these conditions is not exercising "a real choice").

95. That a particular person felt an extreme amount of psychological pressure is of course a claim that can only be verified on a first person basis. Nevertheless, subjective experience is an empirical fact in the sense that the subject had the experience. Nor does saying that subjective states are empirical facts commit us to mind/body dualism or any other metaphysical theory. Indeed, subjective states may one day be completely explainable in terms of neural states, which are verifiable by third parties. *See* RICHARD RORTY, PHILOSOPHY AND THE MIRROR OF NATURE CH. 1 (1979) (discussing the philosophical status of mental states and a skeptical argument that the mind/body problem is itself a non-problem).

96. *See* HEMPL, *supra* note 8, at 22-23 (discussing the requirements of "observation sentences").

candidate upon which to base a claim of duress appears to be unworkable: there does not seem to be some psychological fact by which to distinguish coerced acts from other acts for which an excuse is not available. In addition, there are more general philosophical objections to any theory which attempts to move from factual premises (e.g., regarding the actor's psychological state) to normative conclusions (e.g., whether the actor should be excused for the act). Hume's claim that one can never logically move from an *is* to an *ought* statement is relevant here.[97]

In contrast to conventional, factual theories of duress, moralized theories of duress maintain that claims of coercion 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.[98] In other words, a claim of duress is not simply a legal conclusion drawn from empirical premises concerning the psychological state of the actor; rather, the determination that a particular case is coercive may flow from antecedent moral convictions that the putatively coerced possessed a kind of moral privilege to yield to the threat, or that no person should have to resist a similar threat.

A. PARTIALLY MORALIZED THEORIES

Partially moralized theories of coercion represent a hybrid position between that of traditional and moralized theories of duress. They continue to link coercion to the psychological capacity of the actor to resist a threat, but they also place normative limits on such claims. On one version of this view, the binary notion of voluntariness as a condition that either exists or does not exist is rejected in favor of a view that conceptualizes voluntariness as a kind of empirical commodity that varies by degree. In other words, an act may be said to be more or less voluntary.[99] Alternatively, voluntariness may be conceived as a concept with both empirical and normative components.

Professor Dressler's work on duress represents an example of a partially moralized theory. He maintains 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.[100]

---

97. DAVID HUME, A TREATISE OF HUMAN NATURE 203-04 (Palls Ardal ed., 1972) (1740).

98. WERTHEIMER, *supra* note 3, at 7.

99. Sometimes coerciveness is viewed as being a function of numerous factors that occur in varying degrees and combinations. This permits a ranking of different situations along a continuum. One polar extreme is marked by the voluntariness or freedom of the act and the other by the state of being coerced.

100. Dressler, *supra* note 22, at 1365.

The normative component is evident in Dressler's assessment that duress excuses when the actor lacks "a fair opportunity to avoid acting unlawfully."[101] Yet the concept of coercion is linked to voluntariness insofar as it reflects normal human capabilities to resist a particular threat. The central question that must be answered is whether "we could fairly expect a person of nonsaintly moral strength to resist the threat."[102]

Partially moralized accounts may appear to best serve the legal system's need to set determinate standards of conduct, while recognizing that the underlying justification for duress is related to the actor's inability to resist coercive threats. In both respects, partially moralized theories come closest to the utilitarian theory of duress to be proposed in Part IV of this Article. The problem with these views is that they continue to employ either a gradational or a hybrid concept of voluntariness. The former is conceptually incoherent while the latter is indistinguishable from radically moralized accounts.

As shown previously,[103] gradational accounts of voluntariness are incoherent insofar as they contend that voluntariness is an empirical commodity that exists in greater or lesser degrees. Defenders of this view must be able to specify some psychological entity that increases and decreases depending upon the available external options. Of course, there is no such entity. At best, the gradational view of voluntariness is a metaphorical way of cloaking what are basically normative distinctions in quasi-empirical guise.

Nor does the hybrid notion of voluntariness survive scrutiny. In reality, the idea of voluntariness as a kind of half-empirical, half-normative concept most closely captures and explains our intuitions about voluntariness. Voluntariness is similar to concepts such as health, competence, or justice. Certain factual conditions must be met in order for a particular act to be considered voluntary (i.e., the agent must act in the technical legal sense of the word) but there is also a normative baseline that serves to distinguish the voluntary from the involuntary. Coerced acts are said to be involuntary when they consist of voluntary acts that fall below the baseline by which we judge the normative aspects of conditions precipitating the choice.

The problem with this hybrid notion of voluntariness is not simply that it requires us to abandon the usual legal meaning of the word voluntary—i.e., the sense in which voluntariness simply requires an act in the Austinian sense of the word.[104] The deeper problem is that such views collapse in-

---

101. *Id.*

102. *Id.* at 1367 (emphasis omitted).

103. *See supra* notes 51-56 and accompanying text (discussing a similar problem with the distinction between internal and external causes of behavior).

104. It was Austin, in his lectures on jurisprudence, who first delineated the mental and physical dimensions of an act. In doing so, he followed the assumptions of Cartesian dualism in distinguishing the mental from the physical and in assuming a causal relationship from mental to physical states. JOHN AUSTIN, THE PROVINCE OF JURISPRUDENCE DETERMINED

variably into some version of a radically moralized account. This is because, where voluntariness is viewed as a hybrid of psychological fact and normative judgment, the psychological aspect in each case will always be identical. If the agent has acted in the legal sense of having possessed a mental volition which eventuated in a bodily movement, the distinction between the voluntary and the involuntary, the coerced and the uncoerced, will be wholly normative in nature. In other words, all of the work will be done by the normative half of the hybrid because the empirical component—the volition and the act—will be satisfied in every case where coercion is claimed. Where the act requirement is not met (e.g., where the person has acted while unconscious) the person will be excused for this reason, and not because the act was coerced.

It might be objected to here that what Professor Dressler and others have in mind is neither a gradational nor a hybrid concept of voluntariness. Rather, the empirical and normative aspects of voluntariness may be difficult or impossible to disentangle. The question as to whether a particular act is voluntary requires us, it might be argued, to consider whether a person of nonsaintly moral strength, as Professor Dressler puts it, could have reasonably resisted in a particular situation.[105] This is, at once, an empirical and a normative question. It is empirical in the sense that we think about our own likely conduct: what would we do in such and such a situation? What could be done in any such case? And it is normative in that considerations of what is fair to require of persons enter into the analysis. In this sense, questions of voluntariness are like questions of negligence. They are answered by reference to objective normative standards that reflect our experience about the way in which reasonable persons do, and should, behave.

I suggest that this view is essentially correct as a description of what we do in considering putative cases of duress. Further, it goes a considerable distance in explaining why we excuse in cases of coercion. But this explanation relinquishes any notion of voluntariness consistent with the traditional or will theory of duress. It is fundamentally a normative theory of duress in the sense that there is no specifiable psychological fact that distinguishes the coerced actor from the uncoerced. In this sense, partially moralized theories function in a similar fashion to radically moralized theories, though they retain a theoretical attachment to the concept of voluntariness.

B. RADICALLY MORALIZED THEORIES

Radically moralized or contextualized theories of coercion represent a total rejection not only of traditional accounts of coercion, but of any ac-

(1832). *See* H.L.A. HART, PUNISHMENT AND RESPONSIBILITY 98-99 (1968) (discussing Austin's act requirement).

105. Dressler, *supra* note 22, at 1367.

count which ties the determination of coercion in any given case to some empirical consideration. Such views are "moralized" in the sense that they view coercion to be exclusively a function of considerations such as whether the proposal should be considered a threat, whether the options left to the victim are unfairly limited, and whether we should hold persons in such situations accountable for their acts. In his book, *Coercion*, Professor Wertheimer develops what he describes as a radically moralized account of coercion. While this section will investigate some of the particular features of his theory, many of the comments will apply to moralized accounts of coercion generally.

Wertheimer begins by drawing 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.[106] While these are the three senses in which the term "involuntariness" is used, there are important moral differences among these cases. Most particularly, he suggests that while the first two types of senses represent types of excuses in the legal paradigm, the constrained volition characteristic of duress operates as a type of justification.[107]

Defenders of the traditional theory of duress conceptualize the defense as a type of excuse. On this view, the victim of duress is viewed not to be responsible for the coerced act in that his will has been overcome by external forces. In contrast, Wertheimer views duress as a kind of justification.[108] The coerced act is justified as are acts of necessity or self-defense: the law privileges the actor to act in a manner that would otherwise be illegal on the grounds that the harm with which the actor is threatened is greater than the harm that he causes.[109]

In order for his theory to be consistent with existing law, Wertheimer must meet two related objections here. First, if duress is treated as a type of choice of evils justification, how is it distinguishable from the defense of necessity? Second, how does this account square with the fact that the law often recognizes the defense in situations where the harm produced by the actor equals or exceeds the harm with which he is threatened?[110] Wertheimer responds to these potential problems by arguing that duress is an "agent-relative" justification, whereas necessity operates in an agent-neutral manner.[111] In other words, necessity is permitted as a defense only

106. WERTHEIMER, *supra* note 3, at 9.

107. *See id.* at 166.

108. *See id.* at 29.

109. In the most general sense, acts are excused when the actor is not responsible; they are justified when he is not to be blamed because the act is socially condoned. *See infra* Part IV.B (discussing this issue under the utilitarian account).

110. *See infra* notes 173-76 and accompanying text (discussing the abrogation of the common law rule that barred the use of duress as a defense in murder cases, along with other situations involving great social harms for which duress may provide a defense).

111. Agent-neutral justifications are those the rules of which can be formulated in general

where the harm produced by the act is less than the harm that would result in the absence of the act, where these harms are weighed from a morally neutral vantage point.[112] Duress, Wertheimer maintains, permits the weighing of harms from the particular standpoint of the actor. Thus, the coerced party may consider the special relevance to him of a threat to himself or to a loved one in a way that transcends a mere detached balancing of harms.

The agent-relative account of duress does not withstand critical scrutiny. In essence, this account seeks to privilege certain aspects of the actor's particular circumstances while refusing to privilege still other aspects. It is not clear what principle Wertheimer uses for privileging some but not other aspects of a person's contextualized identity. To refer to an example he uses himself, an actor may be justified in taking the life of another in order to save the life of his wife.[113] Thus, while on an agent-neutral account there might be no way of justifying such an act, on an agent-relative account, the greater interest the actor possesses in protecting the life of a spouse or a loved one takes precedence over the interests of third parties.

What if the actor, it might now be asked, genuinely values the life of her dog over the life of another person? What if she prefers the life of her husband over the existence of our entire civilization? What if the actor has a phobia against even the slightest experience of pain and takes the life of another "under duress" rather than face a kick in the shins? Would any of these be justified on Wertheimer's account?

Wertheimer has responded that only rational agent-relative justifications are privileged.[114] This qualification, however, does little to advance the analysis. The problem here is similar to that faced by his account of voluntariness.[115] If by rational we include all that the agent herself thinks is rational, this subjectifies the standard to the point where anything, given the existence of certain beliefs and feelings, can be justified. If the actor truly loves his wife or his dog, it might be rational for him to sacrifice the greater interests of others in order to preserve his own interests. In sum, a subjectified account of rationality entirely undermines any attempt to place limits on the defense. On the other hand, if there are objective constraints on what is determined to be rational, what are they, and why are only these

---

form without referring to the individual who is the subject of them, while agent-relative justification takes into account the subjective beliefs, values and preferences of the individual. WERTHEIMER, *supra* note 3, at 166. The distinction is discussed in THOMAS NAGEL, THE VIEW FROM NOWHERE 158-59 (1986).

112. The drafters of the Model Penal Code recognize the value of the agent-neutral approach. *See* MODEL PENAL CODE § 3.02(1)(a) (1962) (requiring that the harm which the actor sought to avoid be greater than the harm produced in the act).

113. WERTHEIMER, *supra* note 3, at 107.

114. Professor Wertheimer made this point in conversation at a symposium on Coercion and Exploitation at the University of Denver School of Law, Denver, Co. (Mar. 14-15, 1997).

115. *See supra* notes 69-82 and accompanying text (discussing the internal struggle for actors deciding, voluntarily and involuntarily, to accept offers from others).

privileged? Who decides, moreover, what is rational? More importantly still, privileging only rational agent-relative interests pulls again in the direction of an agent-neutral account and threatens to collapse the distinction between duress and necessity.

It might be said in defense of Professor Wertheimer's view that characterizing duress as a rational agent-relative defense promises the best of both worlds. It is society's way of doing exactly what the criticism says it does: it privileges those, and only those, personal interests that society seeks to protect, and only to the extent that society seeks to protect them. Actors are justified in preferring the life of a loved one over that of a stranger, but may not sacrifice the latter in order to save their dog because society has determined that these interests are special. Even these special interests are limited, however, in the sense that their protection will not justify any harm no matter how great the magnitude. Rationality, in essence, serves to limit the scope of duress as an agent-relative justification to appropriate circumstances.

The problem with this view, however, is that it represents duress as a species of necessity. Once society has determined that certain (personal) interests will be given specified weight in the moral calculus which serves to justify some acts, there does not appear to be much that distinguishes coercion from necessity. The existence of duress as an entirely distinct legal defense with a different mode of explanation vis-à-vis necessity belies Wertheimer's account. To put it differently, if this account were accurate as a descriptive matter, we might imagine that the defense of necessity would have been appropriately enlarged and modified in order to give effect to such personal interests, much as negligence doctrine has evolved to permit modifications of the reasonable prudent person standard of the case of children, the physically handicapped, and other special cases.[116]

Moreover, we must ask why society privileges some of these aspects of personality while not privileging others. Wertheimer's answer—and any answer that sounds in the language of justification—has to say that society has simply made a determination that these interests are important and should be protected as such. But this elides the differences, from a motivational standpoint, between necessity and duress. Duress excuses certain acts not because of some societal determination that certain *interests* are worthy of protection, but because of our collective recognition of the *motivational dimensions* of the coercive situation.[117] In sum, we understand that the fail-

116. In these cases, the objective standard narrows to represent an objective standard for that class of persons. For example, children are held to the standard of a reasonable child of similar age, experience, and education.

117. Part of what renders Wertheimer's account initially attractive is the fact that coerced actors usually do maximize the good or reduce the harm by surrendering to the threatened evil. WERTHEIMER, *supra* note 3; *Cf.* DRESSLER, UNDERSTANDING CRIMINAL LAW, *supra* note 64, at 275 (arguing that not every coerced act involves a lesser-evils situation). But this is because the threat of some greater harm is necessary as a *motivational* factor in compelling the

ure of the law to recognize such interests will not affect the outcome of the coercive situation. We do not provide a defense because the outcome of the act is on balance good, as with any justification; we excuse the coerced because the law can do no good.

We now turn to a question that goes to the very heart of radically moralized theories of duress such as Wertheimer's: If there is, in principle, no empirical basis by which to distinguish genuine coercion from cases which should not be deemed coercive, then can radically moralized theories be said to provide a determinative means for distinguishing the coerced from the uncoerced? Is there any objective criterion for evaluating coercion claims?

Wertheimer argues that coercion occurs when the victim is "faced with a threat or an otherwise immoral proposal (the proposal prong) that limits the choice of the actor to an impermissible range of options (the choice prong)."[118] Both prongs of the analysis must be moralized, i.e., both must be compared to some normative baseline that serves to distinguish permissible offers from impermissible threats, and reasonable ranges of options from those that are normatively unfair or indefensible.[119]

But are the standards that serve to set the baselines in any sense objective or determinate? When we say that a particular person was coerced in a certain situation, are we saying anything about the world at all? Is there any test that serves to answer these questions from the standpoint of the perplexed decision-maker? If the answer to these various questions is "no", then the numerous criteria for theoretical accuracy discussed in Part I must remain unsatisfied.[120] First, there will be no way of measuring the descriptive accuracy of this theory because it does not generate a determinate set of outcomes. Further, if the theory cannot be said to generate a set of determinate outcomes as a descriptive matter, there will be nothing against which to measure our normative intuitions in particular cases. Nor can such a view satisfy the condition of prescriptive efficacy because it does not even purport to tell the judge how to decide in a particular case. Wertheimer's view simply amounts to the claim that *if* we view a particular range of choice options as sufficiently unfair and *if* we determine that the proposal is immoral, then we will decide that the resulting act was coerced.[121] But how do we decide what is sufficiently unfair and immoral? The judge must step into the gap and decide each case by some measure other than that of the the-

---

actor to comply. It is anomalous to view the coerced actor as one who rationally weighs the competing possibilities, as does the person acting from necessity. The coerced actor acts primarily from fear, love, grief, desperation, etc. The production of good or reduction of evil is a byproduct of their act, while it usually may be thought to directly motivate the act in the case of necessity.

118. WERTHEIMER, *supra* note 3, at 30.

119. *See id.* at 30-39 (discussing the moralization of the choice and proposal prongs).

120. *See supra* notes 7-18 and accompanying text.

121. WERTHEIMER, *supra* note 3, at 30-46 (discussing the moralized conditions for duress).

ory itself. Finally, an indeterminate theory may not even be able to tell us why we permit a defense for duress. It does not tell us why society has chosen to privilege the interests it has privileged, nor does it assist us in deciding how future cases involving other similar interests should be addressed. An indeterminate theory cannot answer these questions.

A moralized theory may be able to meet these challenges if it is successful in tying the determination of coercion to something outside itself. In sum, for the theory to be determinative, it must find its basis in something determinate. Does Wertheimer's account meet this challenge? At first, Professor Wertheimer appears to tell us that coercion claims are truth functional or determinate. He states:

> [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.[122]

He then appears to take it back again. In the following sentence he says that he does "not want to put much weight on the term 'truth'."[123]

The radically contextualized character of coercion claims, in Wertheimer's view, is apparent in a subsequent discussion. He imagines the example of a prisoner who confesses to a crime after the police threaten to beat him.[124] He argues that the courts may rightfully exclude the confession on grounds that it was coerced while the prisoner's accomplices may correctly reach the opposite conclusion in deciding that the prisoner did not show sufficient strength in resisting the threats.[125]

How can the opposite responses to the same situation both be correct? How can the courts be accurate in deciding that there was coercion while the accomplices are also right in denying it? Wertheimer responds that the truth conditions for coercion claims may vary with context.[126] By saying that the truth conditions vary with context, however, he does not simply mean that they vary according to situation, i.e., that some situations are coercive while others are not. This is obviously true. Moreover, as his prisoner example makes evident,[127] the context has to do with the subjective impressions of the evaluator. It is in this sense that the context of the criminal justice system and the context represented by the norms that govern the behavior of criminal accomplices may be said to vary. Coercion is evaluated by different measures depending upon the system of accepted practices and values that govern a particular evaluational standpoint.

Wertheimer provides an example of this sense of contextuality. He

122. WERTHEIMER, *supra* note 3, at 184.
123. *Id.* at 184-85.
124. *Id.* at 181.
125. *Id.* at 181-82.
126. *Id.* at 182.
127. WERTHEIMER, *supra* note 3, at 181.

states that our use of the term "cold" may reflect the subject's standpoint.[128] Thus, the person who says that thirty-five degrees is cold when speaking of a January day in Vermont will be wrong insofar as this is actually mild by these standards. But if the same conclusion is reached by someone newly arrived from Miami he will be correct insofar as it will be cold for the Miami speaker.[129]

This response badly misses the mark, however. Wertheimer commits the fallacy of equivocation here. The term "cold" is used in two different senses in this example. The first sense of the word is used to reflect the normal range of temperature in Vermont in January in some statistical sense. The second example refers to the subjective impression of the Miami speaker. To put it differently, it is not that the concept of coldness remains the same while the truth conditions for its application have changed. Rather, the example points up two different things—coldness in the sense that the temperature has fallen below some normal range of temperatures for a given time and place, and coldness in the sense of subjective experience—which happen to be designated by the same term, "cold."

The problem runs deeper still. While it is obviously true that the application of certain terms will vary with context, and while it is also true that the same fact may take on different significance in varying contexts (e.g., the threat of a punch in the nose may be viewed as sufficient to coerce a second grader to turn over his lunch money while the same threat is not sufficient to coerce an actor to murder an innocent third party) this is not Wertheimer's claim. He argues that the *truth conditions* for application of a concept vary contextually.[130] But this is surely mistaken. By varying the truth conditions for the application of a concept we vary the concept itself. It is one thing to say, for example, that what we mean by cold will be any temperature that falls below the average temperature for a given time and place. In this sense, thirty-five degrees will be cold in Miami in August but not in Vermont in January. It is quite another thing to say that coldness will mean any temperature that falls below the average for a given time and place in one context, while saying that the same term will be used to refer to the subjective impression of coldness in some other context. In this latter example, by varying the truth conditions for coldness, we alter the very meaning of the term. If, in the prisoner example, courts and accomplices use differing criteria for determining the coerciveness of the situation, they will be assessing altogether different things.

If, finally, by a contextualized account of coercion claims, it is meant that each observer may decide for herself whether a given case is coercive, then it is simply facetious to speak of there being truth conditions for coercion claims. In this case, the theory of coercion relinquishes any claim to

128. *Id.*
129. WERTHEIMER, *supra* note 3, at 182.
130. *Id.*

normative force and becomes purely descriptive, rather than explanatory or justificatory. Furthermore, if coercion claims are all invariably subjective, it is not even clear why the individual observer need make any determination about the nature of the proposal and range of choice options at all.

The greatest irony of morally contextualized theories of coercion is that they provide no moral guidance at all, but are merely descriptive. They approach the subject from the standpoint of the external viewer, rather than from any internal point of reference. It is for this reason that they fail as theories of duress.

The criticism marshalled here have been directed at a particular form of moralized theories: those that view the moral considerations relevant to coercion claims as contextualized. Contextualized theories reject the idea that these moral considerations are real or objective—that they reflect something outside of themselves. The utilitarian theory, to which we turn next, seeks to provide a determinative answer to coercion questions by linking the question of whether it is appropriate to excuse to considerations of the efficacy of legal deterrents in such cases.

IV. THE UTILITARIAN THEORY OF DURESS

Utilitarianism constitutes one of the two general movements in modern moral theory, along with various forms of deontological thought, holding that moral propositions are quasi-factual in nature, i.e., that they represent something more than the preferences of the speaker, or the practices of a particular culture.[131] In contrast to deontology which, in general terms, focuses upon the inherent rightness or wrongness of particular acts independent of their consequences, utilitarianism's emphasis is upon the consequences of acts as the object of moral assessment.[132] To the extent

131. Both utilitarianism and deontology are moral realist accounts in that each holds that moral propositions can be understood as something more than subjective expressions of the preferences of the speaker or her culture. In the case of utilitarianism, moral propositions have an empirical grounding in consequences relating to the maximization of happiness, pleasure or preferences. Deontological thought, on the other hand, is grounded upon principles of reason. Classical Kantian deontology purports to be purely rationalistic insofar as it holds that moral propositions can be known in an *a priori* manner. *See* JOHN LAWRENCE HILL, THE CASE FOR VEGETARIANISM 2-7 (1996) (comparing moral realism and moral relativism); *id.* at 13-23 (providing an overview of utilitarian thought and objections to it); *infra* notes 118-27 (discussing utilitarian thought). For a recent defense of deontological thought, see generally ALAN GEWIRTH, REASON AND MORALITY (1978) (attempting to generate a deontological moral theory from the logic and from the generic characteristics of action). For the classic deontological statement, see Immanuel Kant, *Grounding for the Metaphysics of Morals*, *in* IMMANUEL KANT, ETHICAL PHILOSOPHY (James W. Ellington trans., 1983) (arguing that morality cannot be predicated upon contingent empirical phenomena).

132. Where the deontolgist will focus upon either the structure of an act or the motive underlying it, utilitarian thought is more "empirical" in that it looks specifically to the real world results of an action. For example, motive is important to the deontologist insofar as it elucidates the nature of the act by reference to the reason underlying it. The same act performed from different motives will merit radically different assessments from the deontologist.

that the good is viewed by utilitarians as being reducible to some empirical or quasi-empirical property such as happiness or pleasure, applied moral thought and, ultimately, social policy, become completely empirical or scientific, assuming the accuracy of the underlying principle of utility.[133]

While modern utilitarian thought can be traced back at least to Hume,[134] it was philosopher and lawyer Jeremy Bentham who fully developed (and coined the name) utilitarianism, and who tirelessly applied utilitarian theory to a broad range of problems in the law.[135] The defining characteristics of classical (Benthamite) utilitarian thought, after its consequentialism, are its appeal to happiness (or pleasure) as the univocal common denominator of the good,[136] and its egalitarian social dimensions, i.e.,

---

For the utilitarian, on the other hand, motive is of secondary importance; it is relevant only insofar as it indicates likely future conduct on the part of the actor, which might be relevant to whether the actor should be punished for the act. *See* JEREMY BENTHAM, AN INTRODUCTION TO THE PRINCIPLES OF MORALS AND LEGISLATION CH. X (1996) (discussing the relevance of motives in utilitarian thought). This emphasis upon the consequences of an act have led some to observe that utilitarianism makes the right derivative of the good, while for the deontologist, right is primary and good is derivative. RAWLS, *supra* note 9, at 25 (arguing that all "teleological" theories, especially utilitarianism, define the good independently of the right).

133. Utilitarian moral argument takes the form of a simple syllogism in which the major premise is some statement of the (moral) principle of utility, where utility is usually defined in some empirically discoverable manner. The minor (empirical) premise specifies some course of action that will result in the consequences with the greatest utility. Thus, the vision of utilitarianism as a "moral science," as in Bentham's idea of the "hedonic calculus," conceives of social and political policy as being conducted by social planners who employ social scientists to determine the likely consequences of various courses of legislative action. In this manner, morality can be reduced to a social science. *See* BENTHAM, *supra* note 132, at CH. IV (discussing the hedonic calculus and its application to social policy).

134. *See* J.J.C. Smart, *Utilitarianism, in* THE ENCYCLOPEDIA OF PHILOSOPHY (Paul Edwards ed., 1969) (discussing the history and critical analysis of utilitarian thought). The author notes that, while Hume is sometimes considered a utilitarian, he uses the term not so much in a prescriptive or even a descriptive way, but in an explanatory way, e.g., to explain why we pick out certain character traits as good. *Id.* at 208. For this reason, the author argues that it is "not advisable" to classify Hume as a normative utilitarian. *Id.*

135. BENTHAM, *supra* note 132, at introduction. Bentham applied utilitarianism theory to such diverse areas as the relevance of the actors state of mind, *id.* at CH. XIII, sec. 11, the problems of justification and excuse, *id.* at CH. XIII, the proportionality of punishment to offenses, *id.* at CH. XIV, the type of punishment warranted for various offenses, *id.* at CH. XV, and a utilitarian analysis of categories of offense, *id.* at CH. XVI.

136. Bentham equates pleasure with happiness. Indeed, according to Bentham, happiness is the ultimate, systematic unifying value. BENTHAM, *supra* note 132, at CH. V (discussing the types of pleasures and pains). This has two implications. First, all values can be reduced to the pain/pleasure continuum. *Id.* at 12. Freedom, friendship, love, fulfillment, passion—each have moral relevance for the utilitarian only insofar as they embody, or conduce to happiness. Second, there is no other competing value. The appeal to happiness defines, justifies, and exhausts the moral universe. *See* HILL, *supra* note 131, at 16-18 (discussing the problems raised by a systemic appeal to happiness, including the problems of comparing competing outcomes in term of happiness, the problem of quantifying happiness and the more fundamental problem that happiness does not appear to exhaust the universe of moral values after all).

it is the sum of *everyone's* happiness that counts in the determination.[137] For example, in a lifeboat situation, the utilitarian forced to sacrifice one passenger might have to choose to throw over the old, rather than the young, on the theory that the young have more potential capacity for experiencing pleasure. Similarly, the utilitarian might choose to save the scientist who is seeking to discover the cure for cancer over others on the ground that the scientist's survival will conduce to the greatest net utility of society generally.[138] Subsequent utilitarians have rejected pure hedonistic utilitarianism in favor of a number of other positions, including Mill's qualitative hedonistic utilitarianism, Moore's ideal utilitarianism and modern forms of preference utilitarianism which underlie much of twentieth century economic thought.[139] All forms of utilitarianism, as the term shall be employed

137. The Greatest Happiness Principle entails that it is everyone's happiness that counts. This leads to radically egalitarian consequences insofar as the principle treats each person's experience of pleasure the same. This is egalitarian in two distinct senses. First, it is irrelevant from the standpoint of utility which acts give the person pleasure—whether it is pushpin or poetry. If one person finds as much (or more) pleasure in watching Jim Carey films as in attending the opera, this will be given equal (or greater) consideration, accordingly. Second, it is egalitarian in that it is irrelevant *who* experiences the pleasure. Wealth, social class or other indicia of socioeconomic success are irrelevant to the issue of utility because one person's happiness is as good as the next person's. The only limitations on this principle are the differences in the subjects' capacities to experience pleasure and the good which the person might in turn have for others.

138. *See infra* Part V.C (discussing the "problem of justice," which is a function of the social aspect of utilitarian theory).

139. John Stuart Mill maintained that there were higher and lower pleasures, thus introducing a qualitative element into the utilitarian calculus. John Stuart Mill, *Utilitarianism, in* ON LIBERTY AND OTHER ESSAYS 138 (John Gray ed., 1991) ("[S]ome *kinds* of pleasure are more desirable and more valuable than others."). Mill made the move in response to the criticism that classical utilitarianism is a base theory because it holds that all pleasures differed only in terms of quantity, and that quantitative judgments could not be made about differing pleasures. He thus arguably introduced an elitist element into the formerly egalitarian approach. The price paid for this move was that Mill could not claim the same quantitative, mathematical, or scientific appeal for utilitarianism. For example, because the different qualitative levels of pleasure are irreducible in that they cannot be cashed out in terms of some more fundamental concept, there is no way of weighing a lesser quantity of a higher level pleasure with a greater quantity of a lower pleasure. This appears to commit Mill to some form of intuitionistic weighing of goods. *See* G.E. MOORE, PRINCIPE ETHICA (1903) (presenting a defense of ideal utilitarianism, the view that goodness and badness inhere in a state of consciousness produced by the consequences of actions, but that this goodness or badness is not exhausted by the pleasantness or unpleasantness of the state of consciousness. Thus knowledge, beauty and other aesthetic qualities possess independent significance). For a classic statement of preference utilitarianism as it influenced economic schools of thought, see LUDWIG VON MISES, HUMAN ACTION 3 (3d ed. 1949).

> Choosing determines all human decisions. In making his choice, man chooses not only between various material things and services. All human values are offered for option. All ends and all means, both material and ideal issues, the sublime and the base, the noble and the ignoble, are ranged in a single row and subjected to a decision that picks out one and sets aside another.

here,[140] nevertheless share the consequentialist view of morality and the emphasis upon the equal interests of all who are capable of experiencing pleasure and pain.[141]

A. THE UTILITARIAN EXCULPATORY PARADIGM

The utilitarian theory of punishment received its earliest and most systematic formulation in the writings of Bentham.[142] Efficient punishment was to be used as an instrument of social engineering to achieve the greatest happiness (i.e., to prevent as much crime with the least punishment necessary to effect the greatest overall balance of happiness). Thus, the negative utility of punishment itself, a function of the pain caused to the criminal in punishing him, must be considered along with the benefits (in terms of crime prevention) that can be achieved through punishment. In this fashion, the utilitarian seeks to answer two questions—whether to punish, and to what extent to punish—by examining the level of crime prevention that would be purchased with a given amount of punishment.[143]

Punishment serves a number of distinct utilitarian functions. The most important of these functions is that of deterrence,[144] both specific and gen-

---

140. Thus, I rule out so-called "egoistic utilitarianism," the doctrine that holds that, for any act, the right thing to do is that which maximizes the subject's own favorable consequences. *See* Smart, *supra* note 134, at 207 (distinguishing egoistic and universalistic utilitarianism).

141. Thus utilitarianism, at least in its classical form, appears to require that the interests of all sentient beings capable of experiencing pain and pleasure be taken into account. This would include non-human animals. *See* Mill, *supra* note 139, at 40-46 (discussing the utilitarian argument for protection of animals). *See generally* PETER SINGER, ANIMAL LIBERATION (1975) (explaining systematically for the first time the implications of utilitarianism for the animal rights cause).

142. *See* BENTHAM, *supra* note 132.

143. The calculus is, of course, more complicated in its application. In calculating the amount of deterrence produced by the threat of a certain punishment, it is necessary to calculate the probability of apprehension, which will factor into the rational criminal's decision concerning whether to commit a crime. *Id.* at 38-41. Even this, however, is a simple calculus compared to more sophisticated attempts by economists, among others, to assess the economics of crime. Other factors include secondary and long-term considerations such as effects on the populace of criminalizing (or de-criminalizing) certain acts. For an interesting example of this, see Lars Ericcson, *Charges Against Prostitution*, 90 ETHICS 335 (1980) (discussing the possibility that de-criminalizing prostitution might lead to diminishing returns by de-valuing the sexual experience). Similarly, utilitarians must look at the effects, in terms of net utility, of publicly adopting utilitarian modes of analysis. *See* Louis Michael Seidman, *Soldiers, Martyrs, and Criminals: Utilitarian Theory and the Problem of Crime Control*, 94 YALE L.J. 315 (1984) (arguing that utilitarianism may undercut the criminal's own sense of individual responsibility). Thus, utilitarian analysis is inherently dynamic and self-reflexive.

144. Bentham places primary emphasis on the role of deterrence. BENTHAM, *supra* note 132, at CH. XIV. Modern utilitarians have registered contemporary uncertainty regarding the extent to which crime is deferrable. *See* Richard B. Brandt, ETHICAL THEORY (1958), *reprinted in* Richard B. Brandt, *The Utilitarian Theory of Criminal Punishment*, *in* MORALITY AND MORAL CONTROVERSIES 410 (John Arthur ed., 1981). Acts committed in the heat of passion, as well as those undertaken without careful consideration of the possibility of punishment, might not be

eral.[145] However, other utilitarian bases for punishment include prevention,[146] and rehabilitation,[147] among others.[148] The utilitarian rationale for punishment contrasts radically with deontological (i.e. retributivist) theories of punishment in placing primary emphasis upon the consequences of punishment, rather than viewing punishment as a necessary social response to the moral blameworthiness of the criminal act.[149]

In keeping with the largely prophylactic view of punishment, the utilitarian theory of excuse holds simply that acts should not be punished when these functions are not furthered, or where the interest furthered in punishing is outweighed by the evil of punishment itself.[150] Thus, while

---

deterrable. Brandt notes that criminologists today tend to view more crime as falling into these categories than Bentham appeared to believe. *Id.* at 414. Nevertheless, deterrence is still the centerpiece of the utilitarian case for punishment. *Id.* at 414-18.

145. Specific and general deterrence are distinguished in that the subject of punishment is specifically deterred from committing future offenses while the population at large is generally deterred as the result of this same punishment of an offender.

146. Prevention differs from specific deterrence in that the offender is physically prevented from committing an act (e.g., by being jailed) rather than being psychologically deterred by threat of punishment. Prevention serves utilitarian goals to the extent that the prisoner would have committed an act but for his imprisonment. Incapacitation, e.g., where a sex offender is chemically castrated to prevent future sexual assaults, may be thought of as a species of prevention, or as an independent rationale for punishment, depending upon whether prevention is limited to physical incarceration.

147. Rehabilitation serves utilitarian goals in the same way that specific deterrence does—by ensuring that criminal acts that otherwise would have occurred do not. The advantage of rehabilitation over deterrence, from the standpoint of utility, is that the latter is effective only as long as there exists some credible threat of punishment, while the former represents some wholesale character change so that compliance with the law is not dependant upon external legal contingencies.

148. Other utilitarian functions include victim's carthorses (the discharge of anti-social feelings which occur with victimization), incapacitation, and education (the character forming role of punishment upon offenders and the general populace). For a discussion of the latter, see Jean Hampton, *The Moral Education Theory of Punishment*, 13 PHIL. & PUB. AFF. 208 (1984) (defending the idea that punishment should be viewed not as a deserved evil, but as a means of correcting the character of the offender).

149. A great deal has been written on the theory of punishment from a deontological standpoint. For a representative sample of the most important such works, see JOEL FEINBERG, *The Classic Debate*, *in* THE PHILOSOPHY OF LAW (Joel Feinberg ed., 1981) (providing an overview of the utilitarian–retributivist debate); *see also*, IMMANUEL KANT, THE METAPHYSICAL ELEMENTS OF JUSTICE (John Ladd trans., 1965) (1797) (presenting the classic deontological statement arguing, among other things, that punishment should never treat the offender as a means to some social end, and thus may not be used primarily as a means of bettering the criminal or deterring criminal conduct); C.S. LEWIS, A CRITIQUE OF THE HUMANITARIAN THEORY OF PUNISHMENT, *reprinted in* MORALITY IN PRACTICE 262 (James P. Sterba ed., 1984) (critiquing therapeutic models of punishment); H.J. McCloskey, *Utilitarian and Retributive Punishment*, *in* UTILITARIANISM AND CRITICAL ESSAYS (Samuel Gorovitz ed., 1971) (arguing that the principles of obligation and punishment are not reducible to utilitarian terms); Herbert Morris, *Persons and Punishment*, 52 THE MONIST 475 (1968) (presenting a modern deontological argument for the right to punish, along with a rejection of therapeutic and utilitarian models of punishment).

150. BENTHAM, *supra* note 132, at 158. Bentham makes clear that punishment itself is to

utilitarian conceptions of justification do not necessarily diverge greatly from traditional retributivist bases for justification,[151] the contrast in theories is evident in their respective views of the function of excuse. On retributivist conceptions of excuse, one is excused simply because one is not morally blameworthy. Conversely, on the utilitarian account, one is excused where the threat of punishment will not deter the actor from committing the act, or where it will not deter others from committing similar acts, or where some other beneficial consequence in terms of crime prevention comes at too great a price, or simply cannot be achieved at all.

For example, Bentham distinguishes between "cases in which punishment is groundless" and "cases in which punishment must be inefficacious," which correspond roughly with the modern distinction between justification and excuse, respectively.[152] This position, characteristic of utilitarian thought, holds simply that we should excuse in cases where punishment would have no causal efficacy in terms of crime prevention, e.g., where it will not deter future criminal acts or rehabilitate the offender. This obviously does not comport with the usual meaning of the idea of excuse, which conveys the notion that the person is not really to blame, rather than that she could not be deterred by punishment. As we shall see, however, there is a close connection between the undeterrability of certain acts and retributivist notions of moral blameworthiness.

B. ACT AND RULE UTILITARIANISM, OR WHY DURESS IS NOT A JUSTIFICATION

Bentham distinguished between necessity and duress, classifying the former as a type of justification and the latter as a species of excuse.[153] In so doing, he distanced himself from those who view duress as a species of justification analogous to necessity.[154] However, Bentham also did not embrace the traditional "overborne will" view of duress: persons who commit

---

be considered an evil visited upon the criminal by society. He writes: "But all punishment is mischief: all punishment in itself is evil. Upon the principle of utility, if it ought to be admitted, it ought only to be admitted in as far as it promises to exclude some greater evil." *Id.*

151. Both utilitarians and deontologists will agree that good acts are justified. Of course, the utilitarian will emphasize the fact that utility is maximized by the act, and that this is what makes the act good, while the deontologist will argue that the act is justified because it is not morally blameworthy. At a deeper level, however, the utilitarian's focus upon the consequences of the act and the deontologist's emphasis upon the motive of the actor may amount to different aspects of the same underlying rationale: it may be that the actor's motive is viewed to be morally blameless because we approve the consequences.

152. Cases in which punishment is groundless include defenses for what today would be called consent and necessity, among others, both of which are viewed as privileges in tort law and forms of justification in criminal law. BENTHAM, *supra* note 132, at 159-60.

153. *Id.* at 160-62. Duress is included among other defenses which are characterized as excuses today. These include Bentham's treatment of infancy, insanity, intoxication and unconsciousness.

154. *See supra* note 22 (citing authors who view duress as a type of justification analogous to necessity).

acts under conditions of duress should remain unpunished because punishment would do no good, not because their act was unfree.[155] The parameters of duress are drawn by reference to considerations of deterrence, among other considerations, not voluntariness.[156]

Those who view duress in justificatory terms do so on the basis of the observation that duress has sometimes been considered to be limited in its application to situations where the harm caused by the victim of duress in succumbing to the threat is less serious than the harm caused if the threat were to be carried out.[157] This is how some explain the common law limitation of duress to cases not involving homicide.[158]

The justificatory theory of duress, however, faces a number of insurmountable, inherent difficulties. First, the justificatory rationale does not explain the preclusion of duress claims from homicide cases. This is because there may be situations in which committing homicide under the threat of duress would be justified, as where the coerced party takes one life to save the lives of two or more other innocent parties. Moreover, duress has never specifically been limited, either in case law or statute, to situations in which the threatened harm outweighs the harm caused in succumbing to the threat.[159]

Additionally, the justificatory theory of duress makes defenses based on duress redundant with respect to claims based upon necessity. The law should not require two separate defenses for what amounts to the same basis for justification. Finally, it has been noted even by a leading defender of the justificatory theory that, if duress were a justification, we should expect that others would have a duty (or at least a privilege) to assist the coerced party in committing the coerced (and justified) act.[160] The law does

---

155. According to Bentham:

> [M]ust necessarily be ineffectual; because the evil which he sets himself about to undergo, in the case of his not engaging in the act, is so great, that the evil denounced in the penal clause, in case of his engaging in it, cannot appear greater. This may happen . . . in the case of a threatened mischief.

BENTHAM, *supra* note 132, at 162.

156. Of course, only voluntary conduct is deterrable, though not all voluntary conduct is deterrable as a practical matter. *See infra* Part IV.D (distinguishing absolutely and conditionally undeterrable acts).

157. This is the rationale for duress given by LaFave and Scott in their text on criminal law. *See* LAFAVE & SCOTT, *supra* note 24, at 433.

158. *See e.g.*, LAFAVE & SCOTT *supra* note 24, at 434-35.

159. Indeed, most statutes distinguish between duress and various forms of justification. *See* MODEL PENAL CODE § 3.02 (providing for the choice of evils justification); *cf. id.* § 2.09 (elaborating the duress defense, where no special requirement is included that the threatened harm must outweigh the harm committed). Most state statutes draw the same distinction. *See, e.g.*, COLO. REV. STAT. ANN. §§ 18-1-708 (West 1988) (duress) and 18-1-702 (choice of evils); HAW. REV. STAT. §§ 702-231 (Michie 1979) (duress) and 703-302 (choice of evils).

160. *See* WERTHEIMER, *supra* note 3, at 168 (reviewing the argument that, if duress is a justification, no person would have the right to resist the coerced act).

not recognize such a privilege, however,[161] and other attempts to explain away this difficulty have proven unsuccessful.[162] For all of these reasons, the justificatory rationale does not comport with the prevailing legal treatment of duress claims.

Conceptualizing duress as a kind of lesser evils justification also leads to inefficiency from the standpoint of broader utilitarian goals. We can see why this is by recurring to the traditional distinction between act and rule utilitarianism. Act utilitarianism holds that, in any given situation, a person should act in a way that maximizes happiness (or, as in cases involving coercion, minimizes pain) *given the consequences that flow from that act in that situation.* Rule utilitarianism, on the other hand, holds that one should follow the rule that tends to maximize happiness in similar cases even if doing so does not maximize happiness in this particular case. Rule utilitarianism has probably garnered more support among utilitarians, notwithstanding the charge that it is a compromised version of the theory, particularly among theorists concerned with issues of consistency, predictability, and workability.[163] Bentham was an act utilitarian while Mill can arguably be characterized as a rule utilitarian. Rule utilitarianism may be defended on the basis of the negative utility brought about, on act utilitarian approaches, by the need to re-evaluate each situation, by inadequate information and by administrative problems created by lack of consistency in similar situations. Thus, following a rule obviates these costs and difficulties. Certainly from a legal standpoint, the need for consistency and predictability of case law argues in favor of a rule-oriented approach.

An act utilitarian version of the duress defense looks similar to a lesser evils justification. It places the emphasis upon the consequences of the coerced act irrespective of the broader policy implications. The act utilitarian view implies that, all other things being equal, the defense should only be permitted where the harm caused by the victim in succumbing to the threat is less than the harm of the threatened act. The problem with this is twofold: First, on this view, the defendant need not be "coerced" at all in order to seek the benefit of the defense. In other words, where the defendant can show that she avoided a greater evil by succumbing to the threat, she should be able to defend her act on the basis of the necessity defense irrespective of whether or not she can be judged to have been coerced.

Second, and more problematic from a utilitarian standpoint, is that there are good reasons for permitting the duress defense even in cases

---

161. *See* LAFAVE & SCOTT, *supra* note 24, at 375.

162. For a critique of Wertheimer's distinction between agent-neutral and agent-relative acts, see *supra* notes 113-17 and accompanying text.

163. *See* J.J.C. SMART, UTILITARIANISM: FOR AND AGAINST 9-12 (J.J.C. Smart & Bernard Williams 1973) (comparing act and rule versions of the theory and defending act (extreme) utilitarianism). *Cf.* H.P. McCloskey, *An Examination of Restricted Utilitarianism, in* UTILITARIANISM WITH CRITICAL ESSAYS, *supra* note 149, at 204-17 (arguing that "restricted" (rule) utilitarianism fares no better than act utilitarianism).

where the harm caused in succumbing to the threat is equal to or greater than the threatened harm. This is because the duress defense serves values that far transcend the harms inherent in the coercive situation. First, the utilitarian must take into account the negative effects of punishing the offender in a case where punishing may have no deterrent effect. The harm in punishing the offender must be weighed against the difference between the (greater) harm caused and the (lesser) harm threatened in the coercive act.

Additionally, there are more generalized interests at stake. Most basically, the quality of life is significantly reduced where we understand that we may be held legally responsible for acts performed in virtue of coercive threats under conditions not of our making. Indeed, in a culture such as ours, which takes seriously the ultimate importance of personal autonomy, nothing could be more offensive than the prospect that a person would be held accountable for acts in which she is forced to act as an instrument of another's will. Nothing could be more destructive to the sense of well-being, security, and dignity of the innocent than to be twice victimized, once by the coercive agent and again by the criminal justice system. In the most significant way, the duress defense represents the utilitarian commitment to principles of liberal autonomy in the sphere of criminal law.[164]

In sum, there are rules for permitting a defense of duress in cases where the coerced act would not be justified. The criminal law's recognition of this, as in cases where the defense is permitted notwithstanding the fact that the gravity of the offense is equal to or greater than the threatened harm, not only undercuts the claim that duress operates as a kind of justification, but evinces the broader utilitarian goals that underlie the defense of duress.

C. THE UTILITARIAN CASE FOR DURESS

As with any matter to be evaluated from a utilitarian standpoint, the utilitarian case for duress is predicated upon a weighing of juxtaposed factors, in this case the benefit of punishing the victim of duress versus the benefit inherent in not punishing. The argument is that punishing the coerced party is pointless in that it furthers no utilitarian purpose. Even worse, however, punishing the coerced results in other avoidable harms.[165]

---

164. *See* HART, *supra* note 1, at 44-45 (arguing in favor of viewing certain excuses as part and parcel of a "choosing system" of criminal law that permits persons to direct their own lives without the fear of punishment for capricious external influences over which we have no control).

165. *See* GLANVILLE WILLIAMS, CRIMINAL LAW, THE GENERAL PART 756 (2d ed. 1961) (defending a utilitarian conception of duress). This approach to duress can be traced to the thought of Thomas Hobbes:

> If a man, by the tenor of present death, be compelled to do an act against the law, he is totally excused because no law can oblige a man to abandon his own preservation. And supposing such a law were obligatory, still a man

The central tenet of the utilitarian theory of duress, with some qualifications to be discussed in the following section, is that the law should permit a defense for duress in cases where the prohibited conduct could not have been deterred. Where the fear of danger created by the coercive agent is substantial and imminent enough that the threat of future punishment is unpersuasive,[166] such punishment is inefficacious, pointless from a utilitarian standpoint. For similar reasons, punishing coerced acts will serve no general deterrence purposes. To the extent that an act cannot be specifically deterred, there a *fortiori* can be no general deterrence for similar acts.

Rehabilitation and prevention are the next two most important utilitarian reasons for punishment. However, they also serve no purpose in cases of duress. Where deterring a coerced act is inefficacious, rehabilitating the actor, or preventing future acts of a similar variety, is simply unnecessary. Put simply, the victim of duress usually poses no greater threat of future misconduct than anyone in the general population. Thus, rehabilitation and prevention serve no utilitarian purposes here.

The utilitarian nature of duress is also evident, when one considers where it does *not* apply. No defense exists where punishment serves some rehabilitative or preventive purpose. The defense is not available, for example, where the victim has recklessly placed himself "in a situation where it is likely that he would be subjected to duress," as the authors of the Model Penal Code have phrased it.[167] While this limitation on the scope of duress can be viewed from a retributivist standpoint as the result of our recognition that the reckless victim is partially to blame, it also makes sense from a utilitarian standpoint. First, where a person has placed himself in such a situation, particularly in the course of other criminal activities, it is relatively more likely that he will do so again in the future.[168] In this case, present conduct is indicative of likely future conduct. Second, even if the coerced act is itself not deterrable, the reckless conduct leading to one's placing oneself in this situation is. The utilitarian treatment of the victim of duress in this context thus parallels its treatment of the chronically accident prone.[169]

---

would reason thus: If I do it not, I die presently; if I do it, I die afterwards; therefore, there is time of life gained.

THOMAS HOBBES, LEVIATHAN Part II, CH.27 (1969) (1651).

166. The actor would just as well take his chances with the criminal justice system, rather than face the more immediate danger, even where he knows that the act he is about to commit is punishable.

167. MODEL PENAL CODE § 2.09(2) (1962).

168. This is true particularly in cases where the person has placed himself in such a situation as the result of a pattern of behavior likely to be repeated. Whether as the result of criminal activity or other acts in which a sufficient degree of danger exists that the person's conduct can be considered reckless, the likelihood of similar activity in the future is certainly greater than that of the general population.

169. *See* HART, *supra* note 1, at 136-57 (arguing that negligent behavior may nevertheless be the basis for criminal liability and rejecting the view that negligence entails inadvertence).

Punishment serves no utilitarian purpose in the case of conduct that cannot be deterred and this section has suggested that the undeterrability of conduct (as the term shall be defined in the next section) is a necessary condition for a claim of duress. Note that undeterrability is not a *sufficient* condition for coercion because some uncoerced acts might be undeterrable. In short, nothing is gained by punishing those who act under conditions of duress.

Taken by themselves, these considerations may leave the utilitarian indifferent as to whether the victim of duress should be punished. Indeed, to the extent that abrogation of the defense might have salutary effects upon crime prevention (e.g., by eliminating the possibility that some criminals might be emboldened to commit crimes with the hope of falsely asserting the defense)[170] the arguments so far presented might not carry the day in favor of permitting the duress defense. The utilitarian theory of duress can only be made convincing by offering some positive reason for permitting the defense in cases of duress. Put slightly differently, one must show that punishing victims of duress would actually subtract from the net utility of the criminal justice system.

There are three such considerations here, two of which we have already discussed in the previous section. First, and most obviously, we must account, in the utilitarian calculus, for the harm caused to the victim of duress in punishing him. The utilitarian cannot remain indifferent regarding the question of punishment since punishing even the guilty counts as a form of negative utility that must be factored into the assessment. This consideration applies across the board in all cases where punishment will have no benefit from a utilitarian standpoint. All other things being equal, this reason alone will tip the scales in favor of the decision to excuse the actor.

A second consideration involves situations such as those discussed in the previous section, where the harm caused by the victim in succumbing to the threat is less serious than the threatened harm. In cases of deterrable conduct, particularly where the threat is made not against the person who is coerced but against an innocent third party whom the victim may have less incentive to protect, the threat of criminal punishment makes it more likely that the coerced party will resist the threat. It is precisely in this situation, where the act is deterrable but the act would be otherwise justifiable, that the law should provide a defense.

For example, imagine a case in which the victim of coercion is faced with the threat that, if he does not assist the threat-maker in robbing a bank, an acquaintance of the victim's will be killed. Permitting the defense in this case allows the victim of duress to succumb to the threat, thereby preventing the greater evil of murder. Without the defense, where the vic-

170. *See id.* at 19-20 (discussing and critiquing this and other possible utilitarian considerations in favor of eliminating exculpatory conditions).

tim knows that he may be punished for his act, his incentive to resist the threat is considerably increased.

The third utilitarian argument in favor of the duress defense was also discussed in the previous section. It is perhaps the most important of the three reasons in that it touches everyone, rather than only those who are victims of coercive threats: a criminal justice system that punishes victims for acts which are neither initiated nor controlled by them is simply barbaric.[171] Everyone suffers insofar as each of us must live with the knowledge that we may be placed in a situation in which we are forced to choose between our lives or well-being, on one hand, and our liberty, on the other.

In sum, the criminal law's recognition of the defense of duress is utility-maximizing (or, more accurately, negative utility-minimizing) by immunizing the person who acts under conditions of coercion from punishment, by protecting third parties who are immediately threatened by the coercive agent, and by safeguarding the tranquillity and autonomy of others in society. Where the undeterrability of coerced acts constitutes the negative side of the utilitarian case for duress in that punishment serves no beneficial function in such cases, these three considerations represent the positive side by demonstrating the dangers inherent in not permitting the defense.

The forgoing considerations not only represent the underlying normative justification for the duress defense, they will also serve to delimit the scope of the defense. Before discussing the application of the defense to real cases, however, one must first explore some of the complications inherent in the concept of deterrence.

D. ABSOLUTE AND CONDITIONAL UNDETERRABILITY

To this point, this Article has presented a rather simplistic picture of the utilitarian case for duress. Complications arise because deterrence is, in one respect, a relative concept. This is because deterrence depends upon both the level of motivation for performing the illegal act and the countervailing penalty in the event one is apprehended. In the case of a coerced act, the more severe and imminent the threat made by the coercing party, the greater the required penalty that is necessary to deter the act. Of course, where the nature of the coercive threat is sufficiently severe, as where the victim's life is in danger, even an equally severe punishment will not deter since the coercive threat is more imminent and very likely more probable. In sum, the victim would sooner take her chances with the law than face immediate death at the hands of the coercive party.

Thus, in some cases, the gravity of the harm threatened by the coercive agent is likely to be so severe that no punishment with any realistic probability of imposition will deter the act. Such acts will be designated as "absolutely undeterrable." In contrast, acts precipitated by less serious

171. *Accord* DRESSLER, *supra* note 22, at 1365 (arguing that duress excuses where the victim's choices are note only hard, but unfair).

threats, or threats that are less imminent, are deterrable, given a serious enough threat of punishment. In the case of deterrable acts, the utilitarian must ask two questions, one empirical and the other normative. First, how much punishment is necessary to deter the act, given the nature of the coercive threat? Second, does this exact too high a price in utilitarian terms? In sum, is punishment efficient?

To take a pointedly draconian example, we might punish littering with a sentence of life in prison. But this would achieve deterrence at too high a price. We would do better, as a society, to endure a certain amount of littering than to imprison those who engage in such acts (assuming that they could only be deterred by such a stiff sentence).

One can thus distinguish, among deterrable acts, between those that can be deterred only at too great a price and those that can be efficiently deterred. The term "conditionally undeterrable" will designate the former class while the latter group will be referred to simply as "deterrable" acts. The touchstone of the utilitarian theory of duress is its recognition of the core psychological reality of coercion (i.e., the idea that no amount of threatened punishment will trump a more imminent threat of harm if that threat is serious enough). *All such absolutely undeterrable acts precipitated by coercive threats should be excused under the defense of duress, irrespective of the offense committed.* On the other hand, the distinction between the conditionally undeterrable, which should similarly be excused, and the deterrable, which should not, will be more nuanced, and will depend upon a weighing of various factors to be discussed shortly.

The level of utilitarian reasoning is distinct with respect to the absolutely and the conditionally undeterrable. The question as to whether an act is absolutely undeterrable is a function of the nature of the severity of the coercive threat independent of any considerations of punishment. To be absolutely undeterrable, the act will generally have to result from a threat that involves death or serious bodily harm to the actor or to a third party with whom the actor has a significant relationship (usually a close family member). Absolutely undeterrable acts are excused unconditionally because punishment can have no beneficial utilitarian consequence. The three considerations discussed in Section C above militate automatically in favor of exculpation because the threat of punishment will have no countervailing offset.

In contrast, where an act is deterrable, the fact finder must engage in a more extensive evaluation. The question here is whether such deterrence is efficient. The line between the conditionally undeterrable act, which will be excused, and other deterrable acts, which will not, is a function of weighing the benefit of the threat of punishment (its propensity to prevent crime) against the unfavorable effects of punishing.

An example will serve to illustrate here. Imagine a youth confronted by members of a street gang who threaten to assault him unless he agrees to assist them in vandalizing a school. Let us assume that the threat is credi-

ble, that every means of escape is foreclosed and that, as the result of the threat and out of fear for his safety, the youth participates. Intuition tells us that, given these facts, the youth should be permitted to claim the excuse.

The utilitarian theory permits such a defense here. Because the threat did not involve death and assuming that it did not involve serious bodily injury, the act is not absolutely undeterrable. Thus, while the youth will not be unconditionally excused, the judge must ask two questions. What amount of punishment would it take, given a realistic probability of apprehension, to overcome the threat of assault? The likely answer is that a great deal would be required in terms of punishment. Accordingly, the second question is whether such punishment is excessive given the nature of the offense. Put in categorical terms, is the amount of punishment necessary to deter the class of acts consisting of school vandalism excessive? Here the judge must weigh the value of deterrence against the three considerations discussed in section C above: the harm of punishment to the coerced party, the harm differential in succumbing versus not succumbing to the threat (i.e., the harm of vandalism versus the harm of the beating), and the general effects on everybody in terms of the negative influence upon our sense of autonomy and well-being. The answer here, I think most will agree, is that the act should be excused as conditionally undeterrable.

It is time now to flesh out the utilitarian theory of duress by addressing a number of long-standing legal issues relevant to the developing law of duress. While the tripartite distinction between absolutely undeterrable, conditionally undeterrable and deterrable acts requires further discussion, we shall see that a great deal will flow from the distinctions drawn here.

E. APPLICATIONS TO THE LAW OF DURESS

The developing law of duress, both through the common law and, more recently, statute, has confronted a number of problems raised by the inherent nature of duress claims. These issues can be divided into three broad areas: issues concerning the nature and sufficiency of the coercive threat, issues concerning the causal relationship between the threat and the offense, and issues related to the types of offenses for which the defense is available. As shown below, the law is in transition on a number of these issues.

1. The Nature and Sufficiency of the Threat

There exist a number of common law and statutory rules that address the question as to when a threat is sufficiently compelling to be considered coercive. At common law, only the threat of death or serious bodily injury was sufficient to make out a claim of duress; simple battery was not sufficient.[172] A number of statutes today similarly limit the scope of applicable

172. *See, e.g.*, D'Aquino v. United States, 192 F.2d 338, 359 (9th Cir. 1951) (holding that

threats;[173] some even stipulate that the coercive demand be limited to the threat of death.[174] More recently, however, the law has recognized less severe threats as coercive. At least one pair of commentators has argued that the requirement of serious bodily injury should be relaxed,[175] while the Model Penal Code does not place any specific restrictions upon the nature of the threat required for duress.[176]

The "person of reasonable firmness" standard embodied within the Model Penal Code permits judges to render more nuanced decisions in situations where, given the less serious nature of an offense, a less serious threat may indeed be subjectively experienced as coercive.[177] Thus, both the Model Penal Code and the utilitarian theory of duress recognize the relativity of coercive threats, i.e., the fact that we might want to encourage succumbing to the threat in one situation but not in another, depending upon the gravity of the offense.

The common law's limitation of the defense to threats involving death or serious bodily harm limited the defense to what we have characterized as absolutely undeterrable acts. The recent transition to the more liberal standard for the type of threats required indicates the law's willingness to consider duress claims in less serious situations—those involving the more highly nuanced reasoning of the conditionally undeterrable. The utilitarian theory supports this more liberal attitude to duress claims.

There are two related issues concerning the law's approach to the question of the sufficiency of threats. First, there is a split of authority regarding whether the threat must involve some imminent danger, or whether a more remote danger is sufficient.[178] Second, there is controversy regarding whether the coerced actor must be personally threatened, or whether the threat of harm to a third party is enough for a duress claim.[179]

---

duress is limited to threats of death or serious bodily harm in the treason case of Tokyo Rose).

173. *See, e.g.*, GA. CODE ANN. § 16-3-26 (West 1998) (requiring death or serious bodily injury); 720 ILL. COMP. STAT. 5/7-11 (West 1996) (same).

174. *See, e.g.*, MINN. STAT. ANN. § 609.08 (West 1987) (requiring death only).

175. *See* LAFAVE & SCOTT, *supra* note 24, at 378.

176. The Model Penal Code provides a defense where a "person of reasonable firmness in his situation would have been unable to resist." MODEL PENAL CODE § 2.09(1) (1985). There is no substantive limitation on the nature of the threat.

177. *See* LAFAVE & SCOTT, *supra* note 24, at 434 (discussing the position of the Model Penal Code and arguing that it represents an objective test that, nevertheless, permits for more nuanced decisions).

178. Most state statutes still require that the threat be imminent. *See, e.g.*, GA. CODE ANN. § 16-3-26 (West 1998); 720 ILL. COMP. STAT. ANN. 5/7-11 (West 1992). Some require an even more direct threat. *See* MINN. STAT. ANN. § 609.08 (West 1987) (requiring threat of "instant" death). The Model Penal Code's "person of reasonable firmness standard" is more broadly worded. MODEL PENAL CODE § 2.09(1) (1985). Presumably, if a threat was serious enough and appeared to be unavoidable, even a less than imminent threat might be sufficient. This standard has begun to influence some cases. *See e.g.*, State v. Toscano, 378 A.2d 755, 765 (N.J. Sup. Ct. 1977) (adopting the MPC formulation).

179. *Compare* GA. CODE ANN. § 16-3-26 (West 1998) (defendant only) and 720 ILL. COMP. STAT. ANN. 5/7-11 (West 1992) (same) *with* KAN. STAT. ANN. § 21-3209 (West 1998) (defendant or a "spouse, parent, child, brother or sister.").

Both issues are essentially also questions concerning the sufficiency of threats.

How imminent must a threat be before it is considered to be coercive? The utilitarian theory requires only that it be imminent enough to create a credible fear of injury.[180] Whether a particular threat meets this standard must remain a matter of case-by-case analysis. What is important to note for present purposes is that, as a threat becomes less imminent, the coerced act becomes more deterrable. This is because the putatively coerced party may possess other options allowing her to avoid performing the act. Thus, even a threat of death or serious bodily harm may be converted from an absolutely undeterrable act to a conditionally undeterrable act as the threat becomes more remote. Similarly, less serious threats may be transformed from conditionally undeterrable to deterrable as the threat becomes more remote.[181]

Should threats against third parties be the basis for the claim of coercion? The Model Penal Code, adopted more recently than many existing state statutes, permits the defense where a third party is threatened.[182] The utilitarian theory of duress provides that the defense should not be limited to situations in which the defendant, rather than a third party, is threatened. Indeed, as we have seen previously,[183] third parties are in special need of protection in situations where the defendant may have no great incentive to protect the third party. Permitting the defense in these situations allows the defendant to succumb to the threat without fear of punishment, rather than risk the safety of the third party. In sum, the law should extend the defense to this situation precisely because the actor might not be sufficiently coerced to act in a situation where we should encourage such an act. It is in such situations that the defenses of duress and necessity overlap.

2. Causation: Subjective or Objective?

A few words are in order regarding two aspects of the nature of the causal connection between the coercive threat and the act. First, the offense must be a product of the threat in the sense that, but for the threat, the actor would not have committed the act in question. Thus, the coercive

---

180. *Compare* United States v. Bailey, 444 U.S. 394 (1980) (requiring that a prisoner who escaped under the threat of death from another prisoner must prove that the threat was imminent).

181. This explains the common law requirement that duress requires an imminent threat. The less imminent the threat, the less likely the act should be viewed as coercive. *See* United States v. Lee, 694 F.2d 649 (11th Cir. 1983) (no duress without *imminent* threat of physical harm); *compare* Esquibel v. State, 576 P. 2d 1129 (N.M. 1978) (threat by prison guards forty-eight to seventy-two hours before escape is imminent enough, as imminence is relative to the circumstances).

182. The Model Penal Code permits a defense where the threat is made against the defendant or "the person of another." MODEL PENAL CODE § 2.09(1) (1985).

183. *See supra* notes 170-71 and accompanying text.

threat should be a necessary condition to the act in the same sense as that which is required in the law of negligence.[184] From a utilitarian standpoint, this ensures that the act truly was undeterrable and obviates the possibility that, independent of some coercive condition, the actor may repeat the crime in the future.

An issue of potentially greater controversy is whether the standard for coerciveness should be judged from some subjective or individualized standpoint, or whether some objective standard should apply. In favor of the subjective standard is the idea that, to the extent that duress is an exculpatory condition, it should be measured from the standpoint of the person excused. In other words, it is not fair to hold one who has been coerced to a standard that lies beyond her reach. Similarly, a more rigorous standard arguably penalizes those of lesser education, understanding, or judgment.

On the other hand, numerous reasons militate in favor of a standardized or objective measure. First, as a practical matter, objective standards reduce the possibility of fraud and obviate the problems of proof inherent in first-person testimonials regarding their motivation. Second, the subjective standard threatens to deteriorate into a blanket excuse for everyone; put differently, how could we formulate a standard by which to judge a person independent of what she did in fact do in the particular situation that is being judged? Third, objective standards provide a goal by which persons measure their own behavior. Similarly, it establishes a rule that will serve to deter acts falling below that particular standard.[185]

The Model Penal Code adopts a standard which is appropriately relativized to the situation while establishing an objective standard with respect to the actor. The MPC requires that acts be measured in accordance with what "the person of reasonable firmness in [the actor's] situation would have been unable to resist."[186] Thus, salient situational factors such as the nature of the threat, the severity of the offense, and related issues are evaluated, while such personal characteristics as education, intelligence, or courageousness are not.

From a utilitarian standpoint, the "person of reasonable firmness" standard can be interpreted in particular cases to further appropriate utili-

184. *See* PROSSER AND KEETON, *supra* note 37, at 263-72 (discussing of the cause-in-fact requirement in tort law). I leave aside here any discussion of whether some analogue of the "substantial factor" test might be required. Vaughn v. Menlove, 132 Eng. Rep. 490 (1837) (providing the classic discussion of the arguments on both sides).

185. *See* PROSSER AND KEETON, *supra* note 37, at 173-75 (discussing the subjective/objective dilemma in negligence law). While this standard is usually viewed to be a function of the "duty" element, it emerges as an aspect of the "causation" requirement in cases involving breach of informed consent, where it must be decided whether the patient's own subjective preferences should be used, or those of some "reasonable person." *See, e.g.*, Scott v. Bradford, 606 P.2d 554 (Okla. 1979) (reviewing the arguments for both sides and adopting the subjective standard in an informed consent action).

186. MODEL PENAL CODE § 2.09(1) (1985).

tarian values and, in general, to strike a suitable balance between deterrence and the protection of individual autonomy. Thus, the utilitarian standard can be appropriately relativized to situations to take into account particular issues raised by the nature of the threat or the act committed. Moreover, the objective standard serves utilitarian administrative values such as consistency of application as well as eliminating the problems of proof inherent in administering a subjective standard.

3. The Nature of the Offense

Yet another issue that remains to be considered is whether duress ought to be available for only certain types of offenses. As discussed above,[187] the excuse was not available as a defense to murder at common law,[188] nor is it permitted under the case law and statutes of numerous jurisdictions today. Additionally, duress is generally not accepted as a defense in two other areas. First, of relatively recent origin, duress is not permitted as a defense to a charge of contempt of court for failure to testify in a criminal proceeding.[189] One recent study found no reported federal case in which the defense was successfully maintained by an intimidated witness.[190] Moreover, most federal circuits have held the defense unavailable as a matter of law in these cases.[191]

Second, duress has rarely been successfully argued as a defense to prison escape.[192] When escape is the product of coercion, the threats involved are typically serious. These cases usually involve the threat of death, severe beating or sexual assault by another prisoner.[193] Thus, they often would qualify as absolutely undeterrable acts as we have defined the term.

Should defendants be precluded from raising duress as a defense in these types of cases? As we have argued previously, the utilitarian theory of duress accords with the more recent trend, as represented by the Model Penal Code, that does not place a restriction upon the types of offenses in

---

187. *See supra* notes 174-76 and accompanying text.

188. *See* WILLIAM BLACKSTONE, COMMENTARIES, Bk. IV § 30 (1765) (stating that duress is never permitted as a defense to murder); *see also* State v. Finnell, 688 P.2d 769, 773-74 (N.M. 1984) (duress is no defense to homicide); Taylor v. State, 130 So. 502, 504 (Miss. 1930) (same).

189. This rule originated in dicta in a footnote in Justice Frankfurter's opinion in *Piemonte v. United States*, 367 U.S. 556, 557 n.1 (1961).

190. WERTHEIMER, *supra* note 3, at 158.

191. *E.g., In re* Farrell, 611 F.2d 923, 924-25 (1st Cir. 1979); United States v. Damiano, 579 F.2d 1001, 1003-04 (6th Cir. 1978); Dupuy v. United States, 518 F.2d 1295, 1295 (9th Cir. 1975); Taylor v. United States, 509 F.2d 1349, 1350 (5th Cir. 1975); United States v. Handler, 476 F.2d 709, 712 (2d Cir. 1973); *In re* Kilgo, 484 F.2d 1215, 1221 (4th Cir. 1973); LaTona v. United States, 449 F.2d 121, 122 (8th Cir. 1971).

192. *See e.g.*, United States v. Bailey, 444 U.S. 394 (1980); Dressler, supra note 21, at 283 (discussing the limitations placed on the defense in the prison escape situation).

193. *See, e.g.*, Debra T. Landis, Annotation, *Propriety and Construction of "Totality of Conditions" Analysis in Federal Courts Consideration of Eighth Amendment Challenge & Prison Conditions*, 85 A.L.R. FED. 750 (1987) (providing an overview of prison conditions generally and the threat of physical and sexual assault).

which it may be raised. As we maintained in Part IV.B, the "no murder" rule cannot be justified on grounds that murder can never be justified, particularly where more lives are at stake if the coerced actor resists. While one could argue that life is an absolute value such that murder can never be justified or excused, this principle is not recognized elsewhere in the law. We do excuse and sometimes justify the taking of innocent life, as when the defendant is insane or is reasonably mistaken in acting in self-defense.[194]

Different issues are raised by the refusal to testify and by jailbreak situations. As a general matter, of course, both represent issues central to the administration of the criminal justice system. In addition, the possibility of deceit and collusion are arguably more likely in these contexts than in others. The question we must ask is whether, given the special issues raised in these contexts, we should treat these types of cases differently from even homicide cases. The conclusion drawn here is that, while these cases involve issues that may make the defense more difficult to claim, there should be no categorical preclusion of the duress defense in these areas.

There are four aspects of the case of the intimidated witness that serve to distinguish it from other cases in which duress is claimed, and which give the state a stronger argument in favor of abrogating the defense here. First, there is the possibility that the criminal justice system will be compromised by permitting witnesses to refuse to testify. This would undermine the value of the law itself by making it more difficult to prosecute and punish those who are in a position to intimidate witnesses (which, in turn, might embolden criminals to commit other offenses with the knowledge that they are immune from punishment). In sum, one might argue that the institutional value at stake is greater even than in homicide cases, insofar as the entire criminal justice system could be undermined.

Second, there is always the possibility that a false duress claim could be used as a cover for bribery and collusion by corrupt witnesses. Third, the law arguably can have a greater deterrent value here than in other situations because the intimidated witness is already in custody. Unlike other contexts, in which apprehension, prosecution and punishment are less likely, the sanction in these cases is imminent: the recalcitrant witness goes immediately to jail for an indefinite period pending their willingness to testify. Finally, the availability of the witness protection program significantly reduces the possibility that those seeking to intimidate witnesses will be able to carry out their threats. Taken together, these arguments create a strong presumption against a claim of duress in such cases.

Even a strong presumption should be rebuttable, however, where these factors are not compelling. The intimidated witness should not be fore-

194. M'Naughten's Case, 8 Eng. Rep. 718 (1843) (concluding that insane delusion was a basis for excuse in the murder of the private secretary to Sir Robert Peel); Shorter v. People, 2 N.Y. 193, 197 (1849) (stating that defendant is justified in taking life of another if he reasonably believed the victim was about to kill her). *See* LAFAVE & SCOTT, *supra* note 24, § 5.7(b), at 456 (discussing reasonably mistaken self-defense).

closed as a matter of law from claiming duress. From a utilitarian standpoint, the claim of the intimidated witness is not likely to be absolutely undeterrable because, even though the threats may be serious, the threatened harm is generally not imminent and the threat itself is arguably avoidable.[195] Nevertheless, the judge or jury should still be required to determine whether, in a given case, a duress claim should stand by weighing the factors relevant to a finding of conditional undeterrability.

The continued integrity and efficacy of the criminal justice system and the arguably greater deterrability of the contempt of court charge will make the defendant's argument of conditional undeterrability more difficult to make. Similarly, questions of the imminence and avoidability of the threat are issues of fact to be evaluated in light of the particular threat, and of the efficacy of the various means of protecting witnesses. Finally, allegations of bribery and collusion can be adjudicated by the finder of fact, as they are in all other contexts. In sum, while duress may be more difficult to maintain, there appears to be no argument for restricting the defense in this context as a matter of law.

A similar approach is recommended in the case of prison escape. The prisoner should not be denied, as a matter of law, the right to raise a legitimate claim of duress. On the other hand, in order successfully to make such a claim, it should be a prerequisite that the escapee has exhausted every remedy open to her, or has a strong reason to believe that such overtures would be ineffective. Additionally, she must immediately turn herself in to the authorities as soon as she is safe from the threat.[196] Where these conditions are met, however, there should arise a presumption that the escape was the product of duress.

F. SUMMARY OF THE THEORY

The utilitarian approach to duress expounded here views duress as an exculpating condition predicated upon the undeterrability of the coerced conduct. Duress should not be limited to acts that would otherwise be justified, as where the actor is motivated by the desire to prevent some greater evil. Justified acts may be defended on the basis of the defense of necessity, independent of considerations of whether the act was coerced. Duress and necessity are overlapping defenses, as where the actor is coerced to perform

---

195. Incidentally, in some cases, there may be no explicit threat at all, only a well-grounded fear on the part of the witness. This raises the question as to whether the intimidated witness can raise a duress claim where no threat has been made. Arguably, there are situations where a witness might reasonably be in fear without having received an explicit threat, as where the witness must testify against an underworld figure connected with the intimidation, abduction or murder of other witnesses.

196. These requirements were outlined in a case permitting two female inmates to plead necessity to a threat of battery and/or sexual assault by other female inmates. The escapees were faced with the alternative, "fuck or fight," and chose to flee instead. People v. Lovercamp, 43 Cal.App.3d 823, 825 (1974).

an act that would be justified even in the absence of duress.

The foregoing has further distinguished two categories of undeterrable acts: those that are absolutely undeterrable, and therefore automatically excused, and those that are conditionally undeterrable and excused only where the deterrence value of punishing is outweighed by those factors militating against punishing the actor. Furthermore, as argued earlier, the standards for absolute and conditional undeterrability must be objective, though each case is considered in light of its own external, situational factors. Thus, the "person of reasonable firmness" standard embodied in the Model Penal Code, or some similar standard, should be utilized. We should understand, however, that, particularly in the case of conditional undeterrability, this standard is really a kind of normative placeholder for the type of utilitarian evaluation defended here, as the "reasonably prudent person" standard is similarly used in utilitarian conceptions of the negligence standard.[197]

Absolute and conditional undeterrability are to be decided as issues of fact. In the case of questions of absolute undeterrability, the jury should determine whether the severity and imminence of the threat caused the defendant to act such that the reasonable probability of even grave punishment would have been insufficient to deter the act. On the other hand, the question of conditional undeterrability will involve a more nuanced weighing of those factors discussed throughout this portion of the Article.

Finally, duress should not be limited in its scope of application either to only the most serious threats or to relatively less serious offenses. While it will be more difficult to satisfy the conditions of the defense in these situations, there will be situations in which a careful utilitarian weighing of the various factors militates in favor of permitting the defense.

V. OBJECTIONS CONSIDERED

The critical response to *Bentham's Introduction to the Theory of Morals and Legislation*, which appeared in 1789, and to the publication of John Stuart Mill's *Utilitarianism* seventy-three years later, was voluminous. Equally compendious have been the counter-responses by defenders of utilitarianism. In the course of the ensuing debate, which continues in the present era,[198] the nature of utilitarian theory has evolved and been refined in a number of fundamental ways.[199] Even with these developments, however, the essen-

197. A Learned Hand opinion is the classic statement on this. United States v. Carroll Towing Co., 159 F. 2d 169 (2d Cir. 1947).

198. *See* J.J.C. SMART & BERNARD WILLIAMS, UTILITARIANISM: FOR AND AGAINST (1973) (providing an excellent overview of the debate on utilitarianism).

199. Perhaps three developments are most significant in recounting the development of utilitarianism from its inception with Bentham. First, modern utilitarians have tended to move from act to rule-based models. Second, they have moved away from hedonistic forms of utilitarianism to intuitionistic and preference-based models. Third, a point which is related to the second development, they have moved away from reductionistic, quanitificational models

tial nature of utilitarian thought remains unchanged in its commitment to a consequentialist, interpersonal and egalitarian ethic.

This part of the Article examines four potential objections to the utilitarian theory of duress. Each of the first three objections is an example, applied here to the problem of duress, of a broader criticism of utilitarian theory generally. The first objection concerns the problem of quantifying utility and the criticism that the utilitarian approach to duress is not really determinate. The second argument maintains that utilitarian principles cannot fully account for society's legitimate interest in punishing those who break the law; thus, punishment may be morally warranted in cases where it serves no utilitarian goal. As applied to duress, it might be claimed that the utilitarian theory of duress is not essentially a reflection of the coerciveness of the situation at all; rather, the determination whether to permit the duress defense in a given case is a function of extrinsic policy concerns.

The third objection holds that utilitarian theory might justify the imposition of unjust or excessive punishment, or might preclude the duress defense in situations where it is warranted. Finally, the fourth objection claims that the utilitarian approach to duress does not comport with the recognized legal understanding of excusing conditions.

A. DETERMINACY AND THE PROBLEM OF QUANTIFICATION

The problem of quantifying utility is essentially two distinct types of argument that are sometimes conflated. Most generally, the objection challenges the claim, made by its defenders,[200] that utilitarianism affords a level of mathematical precision that permits useful comparisons between different sets of consequences. The first dimension of this objection is the fundamentally meta-ethical challenge to the notion that value judgments can be reduced to one normative common denominator (e.g., pleasure). This argument, made by normative pluralists of various stripes, holds that ethical judgments involve a host of competing values (e.g., freedom, love, friendship, happiness, security, compassion, etc.) that are irreducible and incommensurable.[201] The second, more practical argument is that, even if it

---

of utility. We see the first and third developments already in Mill's utilitarianism. *See* MOORE, *supra* note 139; Mill, *Utilitarianism, in* UTILITARIANISM, LIBERTY AND REPRESENTATIVE GOVERNMENT 7 (1910) (arguing that there are differing levels of the quality of pleasure, and that pleasure is not all quantitatively reducible to basic units); J.O. Urmson, *The Interpretation of the Moral Philosophy of J.S. Mill,* 3 PHIL. Q. 33 (1953) (arguing that Mill is best understood as a rule utilitarian). The move to an intuitionistic form of utilitarianism is evident in G.E. Moore. Finally, modern preference utilitarianism is evident in the law and economics movement. *See* RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW 6 (2d ed. 1977) (discussing the notion of utility as maximizing satisfaction of preferences).

200. BENTHAM, *supra* note 132, at 38-41 (discussing the way in which utility is to be calculated according to several dimensions, including its intensity, duration, certainty, fecundity (to produce other pleasure). Bentham's thought is pervaded by the quest for a quasi-mathematical rigor. *Id.*

201. *See* Isaiah Berlin, *John Stuart Mill and the Ends of Life, in* FOUR ESSAYS ON LIBERTY 173-

is conceded that utility is reducible to pleasure or happiness, pleasure and happiness themselves cannot be distilled into discrete units, the relative quantities of which can be measured and compared as against the quantities that might be realized by some other set of consequences. Basically, the utilitarian's asserted pragmatic advantage over competing moral theories—that utilitarianism is more workable, more "scientific," and less fuzzy in its application—is challenged by those who raise the quantificational problem.[202]

In the case at hand, one might argue that the mode of utilitarian balancing described here, particularly in the case of conditionally undeterrable acts, is *ad hoc*. The utilitarian theory requires outweighing the benefits of crime prevention against its costs. This requires balancing, in general terms, the lives, physical well-being, or preserved property interests of those who would have been victimized by a putatively coerced act but for the threat of punishment, against such costs as the harm of punishment and the autonomy of all of us. How can we hope to arrive at anything nearing a mathematically precise solution to these problems, it might be argued, particularly in drawing the line between the conditionally undeterrable and the deterrable?

The problem is not simply the empirical problem of ascertaining whether, and to what extent, crime will be deterred or the sense of autonomy of others will be preserved, etc. Even where we could be certain of the consequences, and of the consequences of alternative courses of action, the problem here involves how to weigh these competing considerations in anything approaching a precise manner.

The utilitarian response to the quantificational problem has been

---

206 (1969) (discussing the relationship between moral pluralism and toleration in Mill's thought, and in liberal thought generally); JOHN GRAY, ISAIAH BERLIN 38-75 (1996) (discussing Berlin's pluralism); JOSEPH RAZ, THE MORALITY OF FREEDOM 395-407 (1986) (attempting to reconcile a soft liberal perfectionism with value pluralism).

Moral pluralism contrasts with moral relativism in that the pluralist is still a moral realist—i.e. the pluralist believes that moral values are real, not simply social constructions. Yet, for the pluralist, there are a diverse number of moral goods irreducible to some common denominator. Thus, moral goods cannot be ordered or prioritized in any rational way. *See* GRAY, *supra* 141-68 (discussing this aspect of pluralism as it relates to justifying liberalism's commitment to freedom).

202. It follows from a commitment to pluralism that the utilitarian quest to define and reduce all ethical values to pleasure is illusory. Consequently, different value options cannot be quantified or measured against one another. Thus, all moral pluralists reject the classical utilitarian reduction of all good to pleasure. John Stuart Mill rejected both the classical utilitarian reduction and the pluralist approach. He is not a pluralist because he still believes that different goods can be rationally ordered. *See* MILL, ON LIBERTY, at CH. II (arguing that those who have experienced two kinds of values can best judge which is qualitatively better). *See e.g.*, J.L. MACKIE, ETHICS: INVENTING RIGHT AND WRONG 125-29 (1977) (discussing the impracticality of quantifying happiness); BERNARD WILLIAMS, MORALITY: AN INTRODUCTION TO ETHICS 92-100 (1972) (discussing a number of problems related to difficulties in quantifying happiness).

largely to accept the objection, to relinquish the more extreme claims to mathematical precision characteristic of hedonistic forms of utilitarian thought, and to maintain that utilitarianism still provides a far greater degree of normative guidance than any competing theory.[203] The quantificational problem militates most forcefully against classical hedonistic forms of utilitarianism; even here, however, the utilitarian can safely maintain that benefits and costs can be weighed against each other in a general, intuitionistic manner. Moreover, modern preference utilitarians will argue that choices can be quantified in a manner approaching mathematical precision, as evidenced by recent economic models of choice.[204]

In the case of duress, the utilitarian theory provides a methodology for evaluating problems raised when deciding whether the defense should apply. It is determinative in telling us that duress should not be limited to acts that would be justified, and further gives us reasons for excusing absolutely undeterrable acts. Even with respect to conditionally undeterrable acts, the theory provides guidance regarding the considerations that must be weighed against each other in reaching a conclusion. Of course, there still remains the difficult task of evaluating factual claims and engaging in the normative balancing of various considerations. But no theory can provide a short-cut to this.

In sum, the quantificational problem raises serious doubts with respect to the pretensions to rigor commonly vouchsafed by old-fashioned utilitarians.[205] But it does not represent a fatal objection to the utilitarian theory of duress. Compared with the vagaries of deontological theory, it remains a much more practical, workable, even empirical methodology for moral decisionmaking and social policy.

B. RETRIBUTIVISM AND THE RIGHT TO PUNISH

A second important objection to the utilitarian theory of punishment involves the contention that recourse to utilitarian principles cannot explain or justify society's legitimate interest in punishing wrongdoers. Most basically, retributivists from Kant through C.S. Lewis have argued that society's right to punish is not exhausted by utilitarian principles.[206] Even if no utilitarian consideration were to be furthered by punishing, society nevertheless may, indeed must, punish the moral wrongdoer.

This objection can be recast with respect to duress in the following manner: It might be argued that the utilitarian theory of duress is really not about coercion at all. At least with respect to the question of conditional

---

203. *See* SMART, *supra* note 163, at 62-67 (discussing utilitarianism and calculating future consequences and the difficulties with this).

204. RICHARD A. POSNER, ECONOMIC ANALYSIS OF LAW (2d ed. 1977).

205. *Compare* BENTHAM, *supra* note 132, at CH. IV, *with* MILTON FRIEDMAN, CAPITALISM AND FREEDOM CH. X (1982).

206. *See* KANT, *supra* note 149; LEWIS, *supra* note 149.

undeterrability, an actor could be excused as coerced in a situation where, as a psychological matter, the putatively coercive influence does not rise to the level of duress, but where not punishing serves overall utility. Put slightly differently, there may be instances in which the actor is indeed morally blameworthy such that society has a right to punish, which nevertheless would be excused on the utilitarian account.

One way to respond to this, of course, is to reject such retributivist objections to the utilitarian theory of punishment generally. The defender of utilitarianism could contend that, where utilitarian principles militate against punishing someone, there simply are no convincing countervailing considerations of the deontological variety in favor of punishment. One might suggest, for example, that retributive principles are inherently fuzzy or vague in their application, that the notion of a right to punish is a vestige of primordial, irrational feelings of vengeance, or that we cannot make sense of what it means to be morally blameworthy in nonconsequentialist terms.[207] While these counter-responses on the part of utilitarians are a mainstay of the on-going debate between retributivists and utilitarians, the response here will take a more limited form: whether or not one accepts the retributive objection to utilitarian theory generally,[208] the objection has little weight when applied to the defense of duress.

In order for this objection to apply to duress, the retributivist would have to demonstrate that there are cases in which the utilitarian would permit the defense to an actor who is morally blameworthy and who the retributivist would not excuse. But there do not appear to be any such cases. The retributivist has no *prima facie* reason for holding the actor responsible in the case of absolutely undeterrable acts. It is difficult to see how even the staunchest retributivist would hold a person morally responsible for acts that could not be deterred by the threat of severe punishment. While resisting such threats might be a heroic feat when accomplished by the courageous or the saintly, failing to do so certainly should not be the predicate for a finding of moral turpitude.

Cases of conditional undeterrability may appear more difficult to respond to because the conduct is, by definition, deterrable. Thus, the actor arguably could have resisted the threat. Even here, however, there is potentially a close correspondence between the outcomes driven by the retributivist idea of moral blameworthiness and those entailed by the utilitarian concept of deterrence. An act will be found to be conditionally undeterrable only when the punishment required to deter it and the societal effects of having a rule punishing such acts are excessive, given the nature of the act. To return to the example of the youth coerced into vandalizing a school, we tend not to hold her responsible where the coercive threat is

207. *See* TED HONDERICH, PUNISHMENT: THE SUPPOSED JUSTIFICATION 35-47 (1969) (providing an overview of criticisms of retributivism).

208. I leave open the possibility that there might exist important nonconsequentialist arguments for punishment.

serious and imminent, even though her act might be deterred by the threat of certain execution. We do not believe an innocent person should be forced to choose between the alternatives of imminent assault and execution. Our preanalytic judgments of moral blameworthiness are mediated through the concept of deterrence. Put differently, we view her not to be morally blameworthy for the same reason we view the deterrent value of execution to be excessive in this case.

This is not to suggest that our preanalytic moral intuitions regarding what it means to be coercive will always be consistent with the distributive results of the utilitarian theory of duress. It does suggest, however, that there is a relatively close correspondence between our sense of being coerced and such values as those which militate against punishment on the utilitarian account. In conclusion, even if the retributivist is correct that utilitarian reasons do not exhaust the possible range of justifications for punishment, this does not undermine the utilitarian theory of duress.

C. THE PROBLEM OF JUSTICE

The previous objection held, in essence, that utilitarian reasons are not *necessary* to justify the imposition of criminal punishment. This objection holds that neither are they *sufficient*. Put differently, where the previous objection held that the utilitarian theory of duress is over-inclusive in that it excuses some who are morally blameworthy, the present objection maintains that the scope of the utilitarian theory of duress may be under-inclusive by failing to excuse some who should not be punished.

The problem of justice, in its various manifestations, is perhaps the most serious objection raised against utilitarian theory generally.[209] The objection holds that the utility-maximizing function might be used to justify violations of individual rights. This objection follows closely the tradition of deontological thought, traceable to Kant, that no individual, even the wrongdoer, should be used as a means to some other end.[210]

As applied to issues of punishment and responsibility, the argument holds that utilitarian considerations such as deterrence, rehabilitation, and the maintenance of overall social order may be used to justify the punishment of the innocent.[211] Similarly, utilitarian principles arguably justify the

---

209. The various versions of the theory hold that utilitarianism justifies the violation of the rights of some if doing so can so increase the net utility of others that, on balance, total utility is increased. As Rawls has put the problem, the utilitarian places the good over the right, thus violating principles of justice. *See* RAWLS, *supra* note 9, at 27-33 (discussing the contrast between utilitarian thought and the idea of justice as fairness). As Rawls notes, in this respect utilitarianism is not an individualistic theory because utility is inter-subjective. *Id.* at 29.

210. KANT, *supra* note 131. For a utilitarian response to the requirement that we never treat others as ends, see P. Nowell-Smith, *Utilitarianism and Treating Others as Ends*, 1 NOUS 81 (1967).

211. For a classic statement of the retributivist objection to the utilitarian theory of punishment, see H.J. McCloskey, *Utilitarian and Retributive Punishment*, *in* UTILITARIANISM WITH

imposition of excessive sentences where this may be important in furthering utilitarian objectives such as deterrence. Consequently, the individual should be punished, according to the retributivist, only for her moral wrongdoing, and only to the extent warranted by the wrongdoing. Persons should not be punished for purposes of furthering other social objectives; moreover, the punishment should fit the crime in terms of its severity.

As with the previous objection, however, it is not clear that there is a systematic divergence between the outcomes generated by the utilitarian and retributivist theories. For this objection to have any merit in the context of duress, there would have to be cases in which the utilitarian would punish acts which the retributivist would excuse. Since the utilitarian excuses the absolutely undeterrable, the retributivist is limited to an attack on the treatment of deterrable acts. Yet the retributivist is correspondingly less likely to find coercion where the act is deterrable. Consequently, it is simply not clear that there exist identifiable cases where the utilitarian theory would fail to excuse the morally blameless.

One persistent problem in comparing retributivist and utilitarian theories with respect to the results that each generates is that retributivism does not appear to embody a well-delineated set of principles for determining issues of punishment and excuse. Retributivist thinking is notoriously difficult to evaluate precisely because its underlying notion of moral blameworthiness is not clear. For this reason, one cannot make substantive comparisons with utilitarianism in concrete cases.

Finally, even if the applications of the theory could be compared, and even if results generated by utilitarian and retributivist conceptions do vary in the context of duress, it does not follow that it is the utilitarian theory that must yield. For this objection to have any force, the retributivist would have the burden of demonstrating that nonconsequentialist thought provides a more satisfactory basis than utilitarianism for deciding questions of punishment and excuse.

D. THE UTILITARIAN THEORY OF DURESS AS AN EXCUSE

The last set of objections concerns problems related to the characterization of duress as a type of excuse. There are two related objections here. First, it is arguable that the utilitarian approach to duress does not fit properly into the two-pronged exculpatory paradigm we surveyed in Part II.A. That portion of this Article argued that, in order to be considered an excuse, a defense had to have its basis in some cognitive or volitional defect.[212] While duress is usually thought of as resulting in a volitional defect, particularly on the traditional model, this Article has argued here that the exculpatory nature of the defense is not a result of the lack of volun-

---

CRITICAL ESSAYS, *supra* note 149 at 361-75 (rejecting utilitarian attempts to rehabilitate their position against the claim that it leads to unjust punishment).

212. See *supra* Part II.A.

tariness. Thus, it is not clear on what basis duress should be categorized as a type of excuse.

This objection is not fatal to the utilitarian approach because, even if the argument is admitted, it does not go to the inadequacy of the utilitarian theory itself, but simply to its characterization as a type of excuse. Nevertheless, the argument will not be conceded here because duress does act as a kind of excuse in the utilitarian paradigm. It acts as an excuse in the sense that it goes to the condition of the actor, rather than the act, thus distinguishing it from a justification. Moreover, though the condition of the coerced actor is not characterized by an inability to act volitionally, the act is against the will of the actor in the sense that it results from the coercive threat where no countervailing consideration, including the possibility of serious punishment, can be effective in changing the course of action. This is all that should be required to excuse an act on the utilitarian, or any other, paradigm.

A second possible objection holds that, on a utilitarian account, duress cannot be distinguished from other excuses. To the extent that the touchstone of the utilitarian theory of duress is the undeterrability of certain conduct, the same thing may be said of other excuses. The insane actor or one who acts from provocation will be excused only after a similar weighing of deterrence versus the harms of punishment. Thus, duress cannot really be distinguished from these other defenses in categorical terms. What this arguably indicated is that these various legal defenses cannot be understood in utilitarian terms, as utilitarianism has no basis for distinguishing among them. Put in terms of the criteria for theoretical adequacy discussed above,[213] it might be said that the utilitarian theory lacks descriptive accuracy insofar as the theory cannot distinguish among various excuses which are recognized as distinct in the law. It further lacks prescriptive power insofar as utilitarianism can provide no reason to account for the recognized taxonomy of legal excuses.

The response here may be put simply: the utilitarian approach can distinguish the various legal excuses in two ways. First, it can distinguish on the basis of the condition which gives rise to the undeterrable conduct. Insanity, provocation, mistake, duress, and any other excuse can be delineated with respect to the factors which have resulted in the act. Each such condition will involve different issues of proof and various related legal questions (e.g., what conditions should be required as a matter of law to be able to state each defense) which alone give sufficient warrant for delineating these various excuses.

Secondly, various defenses will warrant different analysis and treatment. For example, while insanity will be excused in virtue of the same undeterrability as is duress, different issues arise with each defense. In the case of insanity, the actor remains a threat to society in a way that the co-

213. *See supra* Part I.

erced actor does not. Thus, utility will be maximized only by prevention and treatment in a way that is not required for the victim of duress. Similarly, while the provoked act might, in extreme cases, be undeterrable, it is arguable that the actor remains a threat to society in a way that the coerced actor does not. (If the actor is easily provoked, or if such instances are likely to repeat themselves, there may be significant utilitarian reasons for refusing to exculpate the actor.)

In conclusion, the utilitarian approach to duress is consistent with both our prevailing conception of exculpation, and with the distinction between the various legally recognized excuses.

VI. CONCLUSION

The utilitarian theory of duress is superior to both traditional and moralized accounts of duress with respect to the five criteria for evaluating theories discussed in Part I. Moralized accounts, particularly radically moralized theories such as that of Professor Wertheimer, cannot be evaluated on most of these criteria at all. This is because the moralized account does not purport to tell us what a judge should do in a given case.

Traditional accounts of coercion, on the other hand, do seek to provide a means for evaluating actual cases. They do so unsuccessfully, however. The problem is that they depend upon a concept of voluntariness that either cannot explain the normative force of duress claims or that implicitly appears to be a normative element that undermines the theory altogether.

By contrast, the utilitarian approach to duress provides a high degree of descriptive accuracy. It explains a number of modern trends that reflect the expanded availability of the defense in situations where it would not have been permitted at common law. It is also highly consonant with our pre-analytic normative intuitions in reflecting the relative nature of duress claims. The utilitarian approach vindicates modern sensibilities that would extend the defense to situations involving less serious threats and to those in which the offense itself may be quite serious. Finally, and perhaps most importantly, the utilitarian theory gives us a degree of prescriptive power that cannot be achieved using the other two theories. It explains why we excuse persons who act under coercive conditions. In locating the exculpatory nature of duress in the undeterrability of the act, we also explain why coerced acts are viewed to be against the victim's will. Where the victim acts in a manner that is otherwise foreign to his nature and disposition, the behavior is, indeed, against his will in a very real sense. In this way, the utilitarian approach vindicates our pre-analytic intuition that the coerced act is not freely chosen.