UNC SCHOOL OF LAW

NORTH CAROLINA LAW REVIEW

Volume 70 | Number 5 Article 4

6-1-1992

The Place of Formalism in Legal Theory

Michael Corrado

ESSAY

THE PLACE OF FORMALISM IN LEGAL THEORY

MICHAEL CORRADO*

I understand formalism to be the claim that law is an autonomous area of knowledge. To say that it is autonomous is to say that it is self-contained, that it is not dependent on other areas of knowledge like morality or politics or sociology. In particular it is to say that law is not *reducible* to those other areas, in a sense of reducibility I will try to make clear.

There is a version of formalism that is true. This version of formalism does not claim that law is unrelated to other areas of knowledge. It admits that legal argument makes use of moral, political, and sociological information. But it insists that, although much that is extra-legal is taken into account in deciding a question of law, it remains extra-legal; it does not become part of the law.[1]

I. THE AUTONOMY OF LAW

Interest in logic as a philosophical instrument has taken two different and mutually exclusive forms over the course of the twentieth century. It has found expression, on the one hand, in the attempt to use the machinery of symbolic logic to show that all knowledge can be "reduced" to a *basis*—some small set of elementary or primitive terms and propositions involving those terms—a project known as reductionism. The manifestations of this attempt are logical empiricism, logical behaviorism, logicism in mathematics, and emotivism and naturalism in ethics. Each of these, if successful, would contribute to the larger goal of universal reduction of all knowledge. If the reduction is epistemological

* Associate Professor of Law, University of North Carolina at Chapel Hill. B.A. 1965, B.S. 1966, Pennsylvania State University; A.M. 1968, Ph.D 1970, Brown University; J.D. 1984, University of Chicago.

1. Ernest Weinrib defended a version of formalism in *Legal Formalism: On the Immanent Rationality of the Law*, 97 YALE L.J. 949 (1988). In addition to coming out of a philosophical tradition different from the one in which this Essay is based, Weinrib's position differs in significant ways from the one presented here. As I understand him, Weinrib believes that the internal logic of the law determines what the law ought to be, or how legal terms ought to apply to the facts. *See id.* at 1012-14.

in nature, meant to duplicate the structure of knowledge, then the terms of the basis will be phenomenal or sense-data predicates.

The other form this interest in logic has taken is the attempt to formalize and systematize autonomous areas of knowledge, without reducing them to one another. For each such area the aim is to provide a set of primitive terms and primitive propositions which would generate all the relevant terms and conceptual truths of that area, and which could be applied to the available factual information in that area.[2] The classic (but incomplete) example of such a systematization is Euclid's geometry, and in the last century a number of areas of mathematics were so formalized.

In this century philosophers have applied the method to discrete areas of philosophical interest, and there are studies of the internal logic of the modalities (necessity, possibility, impossibility); of obligations; of knowledge and belief; of semantic relations; of evidentiary relations; and so on. Take, for example, the modalities: necessity, possibility, impossibility. Philosophers have shown not only how to define "possibility" and "impossibility" in terms of necessity, but also how to define implicitly different senses of "necessity" by means of axiom sets.[3] (Such a study might just as easily start with possibility or even impossibility as primitive; but whatever the primitive term is will have to be characterized by

---

2. A primitive term is an undefined term. A primitive proposition is a proposition from which other propositions in the relevant area of study are derivable, and which is not itself derivable. The various terms of geometry are defined, ultimately, by means of certain primitive terms that are not themselves defined. To suppose every term in the system could be defined would be to presuppose a certain circularity; there must be a stopping point to definition. This point of view supposes that something similar is true of the propositions that delineate the conceptual part of a discipline as well. Not every proposition in geometry can be proved; we must start with unprovable axioms and postulates.

3. A nice, informal introduction to the logic of the modalities can be found in ARTHUR N. PRIOR, TIME AND MODALITY (1957). The syntax and semantics of the various systems of "modal logic" have been worked out in some detail. Although the study of formal modality dates back at least to the Stoics, *see* BENSON MATES, STOIC LOGIC (2d prtg. 1961), much of the recent work had its origin in C.I. Lewis's attempt to work out the logic of counterfactual conditionals. One of the simpler systems of logic he worked with had axioms and rules that roughly correspond to the following:

> Proposition One. If something is necessarily true, then it is true in fact.
>
> Proposition Two. If a second thing necessarily follows from a first, and the first thing is necessarily true, then the second thing is necessarily true.
>
> Rule of Inference. If something is a true proposition of this system, then it is a necessarily true proposition of this system.

A number of stronger systems can be built using this as a base. Where *necessity* is taken as primitive, a proposition is *impossible* if it is necessarily false, and a proposition is *possibly true* if it is not necessarily false. *See* CLARENCE I. LEWIS & COOPER H. LANGFORD, SYMBOLIC LOGIC 122-98 (2d ed. Dover 1959) (1932).

means of axioms.) The definitions and axioms are "internal" to the study of modalities; they characterize the meaning of (or at least the structure of the proper use of) the terms. They do not, however, determine the application of the modalities in the real world.

The question of how the modalities are applied in the world—what things are really necessary and which are merely possible and so on—is the subject of a different sort of study. The principles that relate modal terms to nonmodal terms ("If *A* causes *B*, then the proposition that things like *A* are followed by things like *B* is necessarily true"[4]) may be called "bridge" statements. Similarly, in moral philosophy statements relating moral terms to one another and characterizing the use of moral terms by means of axioms are internal statements, and are sometimes called metaethical principles. They are supposed to be conceptual truths—true in virtue of the meanings of the words involved. But the bridge statements of ethics—the claims of act-utilitarianism that rightness and wrongness are determined by happiness and suffering, for example—are substantive principles and cannot be determined just by studying the meanings of the words.[5] Or so at least a "formalism" in ethics would claim.

I will preface the discussion that follows by saying that although the attempt at universal reduction has been convincingly shown to be impossible, the attempt to formalize autonomous areas of knowledge is still very much alive. Though only a small minority of professional philosophers cultivate it in symbolic form, it is informally practiced by many.[6]

---

4. Notice that Hume would reject this claim. For Hume, as he is usually understood, causation is a contingent relation between events. That is, even in a case in which one event follows another and we are inclined to say the first causes the second, the first does not necessitate the second; it might have been otherwise.

5. There is an informal discussion of the distinction in Alvin I. Goldman, *What is Justified Belief?, in* EMPIRICAL KNOWLEDGE: READINGS IN CONTEMPORARY EPISTEMOLOGY 171, 171-73 (Paul K. Moser ed., 1986).

6. William Parent, for example, wrote at the beginning of an article on privacy:

> Defining privacy requires a familiarity with its ordinary usage, of course, but this is not enough since our common ways of talking and using language are riddled with inconsistencies, ambiguities, and paradoxes. What we need is a definition which is by and large consistent with ordinary language, so that capable speakers of English will not be genuinely surprised that the term "privacy" should be defined in this way, but which also enables us to talk consistently, clearly, and precisely about the family of concepts to which privacy belongs.

W.A. Parent, *Privacy, Morality, and the Law*, 12 PHIL. & AFF. 269, 269 (1983). In this passage we can find attitudes common to this sort of philosophy: rejection of a blind reliance on ordinary language, the notion of a family of related concepts, and the emphasis on clarity. The most brilliant and inspiring practitioner of the formal approach may have been A.N. Prior. *See supra* note 3.

A. Universal Reduction

In *The Logical Structure of the World*,[7] Rudolf Carnap gave a very precise meaning to empiricism and then tried to show that empiricism as he defined it is true. He proposed to show that all meaningful sentences could be reduced to primitive propositions involving only logical constants and predicates of sensory awareness (like "appears red" or "has a square appearance"). Reduction was to be a kind of translation; if a sentence involving claims about one sort of thing could be translated into a sentence or set of sentences about other sorts of things, then talk about the first sort of thing would be said to be reducible to talk about the second sort of thing.[8] If such translation were possible for every sentence about entities of a certain sort, then we could do without talk about those entities entirely. Moreover, we could also retain words for the first sort of entity once the reduction was complete; to talk about entities of the first sort would be perceived as simply shorthand for talk about entities of the second sort.[9]

To give a common-sense sort of example, the fact that we talk about "the average person" would lead to some odd consequences if we did not realize that such talk was just shorthand for talk about the totality of persons. To explain it to someone who did not understand it, we might offer a "reduction schema": to say that the average person has a certain property to degree *F* is to say that if we added up the amount of that

---

7. RUDOLF CARNAP, THE LOGICAL STRUCTURE OF THE WORLD & PSEUDOPROBLEMS IN PHILOSOPHY (Rolf A. George trans., University of Cal. Press 1967) (1928).

8. "A concept is said to be reducible to others, if all statements about it can be transformed into statements about these other concepts; the general rule for this transformation of statements for a given concept is called the *construction* of the concept." *Id.* at 15. Carnap did not suppose that the possibility of logical construction answered any "metaphysical" questions at all. The fact that all talk about human beings could be reduced to talk about physical objects, which could further be reduced to talk about elementary experiences, did not mean that human beings did not exist.

9. This is a loose, but I hope not inaccurate, paraphrase of what Carnap said he was doing. Here is his description of the move from the basic experiences ("elementary experiences") to the logical construction of physical objects:

> To begin with, we shall discuss the method of constructing three-dimensional, *physical space* . . . , and then we shall carry out this construction as well as the construction of the visual things which depend upon it . . . . For the constructional system, the most important visual thing is *my body* . . . . It will help us to give definite descriptions of the various senses, so that with its aid we can supplement the domain of the autopsychological . . . . Then we shall describe the construction of the *world of perception* . . . as well as the construction of the *world of physics* . . . , which is quite different from the former. Finally we shall discuss some physical objects ([*e.g.*,] persons . . .), which are required for the subsequent construction of the heteropsychological objects.

Id. at 191. Heteropsychological objects, which are to be constructed logically on the basis of physical objects, include an intersubjective world, cultural objects, and values. *Id.* at 214.

property that all persons have, and then divided the total by the total number of persons, we would get *F*. But there is nothing wrong with retaining the locution "the average person" as a kind of shorthand. Similarly, if the project sketched out in *The Logical Structure of the World* succeeded, we could see all our talk about the world and the things in it as shorthand for more complicated talk about sense-data.

The model for this sort of reduction was the enormous enterprise undertaken by Russell and Whitehead of reducing mathematics to logic, an enterprise known as "logicism."[10] Logicism had been entertained as a possibility for centuries; what Russell and Whitehead undertook to do was to use the machinery of modern symbolic logic to show that one area of mathematics after another could be reduced to logic. What that meant in practice was that they provided definitions of the terms of mathematics in purely logical terms, and then attempted to show that all the truths of mathematics could be derived from truths of logic.[11]

Carnap's aim was to carry this program through the entire body of human knowledge, adding to the logical terms only sense-data predicates. The world was to be structured into sets of propositions of varying levels of complexity: statements of social-cultural fact were to be reducible to statements about the mental events of human beings; statements about the mental events of human beings were to be reduced to statements about physical entities; and statements about physical objects were to be reduced to statements about sensory information.

Carnap's empiricism failed, which is to say, for example, that sentences about human beings could not be reduced to sentences about physical objects without remainder—something irreducibly human would always be left over. This discovery was made possible by the precision with which Carnap and others formulated the thesis of reducibility, and it is one of the things about which it is possible to say that philosophy has reached a conclusion. There are two ways to respond to this discovery.[12] The philosopher could deny any significance to the bit

10. *See* BERTRAND RUSSELL & ALFRED N. WHITEHEAD, PRINCIPIA MATHEMATICA (1910).

11. Here is a simple example: To say that there are three pigs, it suffices to say that there is something *X* which is a pig, and something *Y* which is a pig, and something *W* which is a pig, and that *X* is not identical with *Y* nor with *W*, nor is *Y* identical with *W*. The postulation of the entity *X* and the others is a logical notion as is identity in the logical system of Russell and Whitehead, and indeed in all contemporary logical systems. To say that three pigs and two geese make five farm animals, if pigs and geese are farm animals, requires us to say that *if* there are three pigs (translated as above) and there are two geese (translated in a similar way), then if pigs and geese are farm animals there are five farm animals (translated in a similar way). And this last is a truth of logic.

12. I do not share the trendy view that philosophy does not progress, but only "changes the subject" from time to time. It is the achievement of Carnap and his colleagues to have set

left over; that was Quine's way in *Word and Object*.[13] Or the philosopher could come to the conclusion that biology and psychology are autonomous and irreducible areas of knowledge, and set out to uncover the internal logic of these areas.

B. The Autonomy of the Law

This sketch makes a fairly precise definition of autonomy possible: An area *A* is *reducible* to an area *B* if the terms of *A* can be defined entirely by means of the terms of *B* (along with purely logical terms), and if, with terms so defined, the conceptual truths of *A* turn out to be truths of *B*. In the context of reduction, terms which cannot be further reduced are called primitive terms. An area *A* is *autonomous* with respect to *B* if it is not reducible to *B*. Every autonomous area will have its own primitive terms. If biology is autonomous with respect to physics, then there will be some biological primitives; not everything true of biology will be translatable into purely physical terms.

Morality is an interesting case. From the outset Carnap and others conceded that the sentences used in moral discourse could not be reduced to sentences involving only logical and phenomenal predicates. Rather than take morality as a counterexample to the feasibility of reduction, emotivists declared it to be without meaning: the sentences of morality were said not to convey any knowledge, but really to express emotions the way the expression "Hurrah!" might be used to express an emotion. Those who believe on the contrary that moral discourse does

---

the stage for the dramatic discovery that philosophical behaviorism and phenomenalism could not work, just as Russell and Whitehead showed that a simple and straightforward sort of logicism could not work.

13. But there remains a thesis of Brentano's . . . that is directly relevant to our emerging doubts over the propositional attitudes [*e.g.*, "John believes that . . ."] and other intentional locutions. It is roughly that there is no breaking out of the intentional vocabulary by explaining its members in other terms. . . . Even indirect quotation, for all its tameness in comparison with other idioms of propositional attitude, and for all its concern with overt speech behavior, seems insusceptible to general reduction to behavioral terms . . . . And when we turn to belief sentences the difficulty is doubled. . . .

. . . .

One may accept the Brentano thesis either as showing the indispensability of intentional idioms and the importance of an autonomous science of . . . [statements about the mental], or as showing the baselessness of intentional idioms and the emptiness of [such a science]. . . .

. . . If we are limning the true and ultimate structure of reality, the canonical scheme for us is the austere scheme that knows no quotation but direct quotation and no propositional attitudes but only the physical constitution and behavior of organisms.

WILLARD V.O. QUINE, WORD AND OBJECT 220-21 (1960).

convey information take its irreducibility as signifying its autonomy, and (if they are so inclined) might press on to formulate a set of primitives that would permit the definition of the important terms of morality.[14]

How does all this relate to law? The claim of the modest version of formalism I am advocating is that legal discourse forms an autonomous area of knowledge, and that the sentences or propositions of the law cannot be reduced to sentences or propositions of any other area at all. What it does *not* claim is that legal propositions are unrelated to propositions of other sorts, or that information from other fields is irrelevant to legal conclusions.

Consider the analogy with moral discourse. To say that morality is autonomous is to say that statements about right and wrong are not reducible to, say, statements about people suffering; to say that killing is wrong doesn't just amount to saying that killing makes people suffer. At the same time, suffering is not unrelated to judgments about right and wrong: it is wrong to make people suffer without justification. Without making use of the controversial distinction between analytic and synthetic statements, we may agree that moral terms bear a type of relationship to each other that they do not bear to nonmoral terms.[15] Statements about what actions are morally required may be replaced by statements about what actions it is morally wrong not to do, for example, without any loss in significance; and the statement, "What is morally required it would be morally wrong not to do," we may call an internal statement in moral discourse: it defines one ethical term by means of another. On the other hand, the claim that creating suffering is morally wrong is not of that sort. We may call it a bridge statement: it connects a moral term with a term descriptive of human feelings.

The internal statements of morality are therefore the statements that relate moral notions one to the other and those that delineate the logic of moral notions. The bridge statements of morality are those that relate

14. There are three positions a nonreductionist might take on this further aim of systematization: (1) You might optimistically look forward to the systematic analysis of every autonomous field, with primitive terms and axiomatic propositions, on the pattern of geometry; (2) You might argue that such an aim is mistaken in areas other than logic, and is a waste of time; or (3) You might see such work as productive, even if it does not succeed in the end. Just as Carnap furthered knowledge in his unsuccessful attempt to reduce all knowledge to experience statements, so the unsuccessful attempt at the rigorous formulation of autonomous areas will lead to deeper knowledge of those areas. For example, consider the work of the last 25 years on deontic logic (the logic of obligation statements) and the logic of knowledge and belief.

15. If we were to use this controversial distinction, of course, we might say that the internal statements of an area are analytically true (true in virtue of the meanings of their terms) and that bridge statements are synthetic and a priori. Doubts about the distinction are expressed in WILLARD V.O. QUINE, *Two Dogmas of Empiricism*, *in* FROM A LOGICAL POINT OF VIEW 20, 20-37 (2d ed. rev. 1980).

moral notions to nonmoral notions. To say that morality is autonomous is not to say that moral claims are unrelated to physical, or psychological, or sociological claims. It *is* to say that no set of bridge statements will provide a *definition* for any moral term, and that moral terms are not, therefore, reducible to nonmoral terms.[16]

Similarly, we may distinguish between statements internal to the law and bridge statements. An example of the first might be the statement that there is no criminal liability for actions that are legally justified. This statement is conceptually true—true in virtue of the very meaning of the terms involved in it. An example of the second might be that an action is legally justified if it was carried out in self-defense.[17] The truth of this statement, in those jurisdictions in which it is true, does not follow from the meaning of the terms involved. The challenge is to show that these are in fact different sorts of statements in the law, that legal notions cannot be defined by way of nonlegal terms in bridge statements, and that therefore the law is irreducible and autonomous in the same sense in which morality is autonomous. I will consider just two sorts of reduction here, but I think my argument can be extended to any proposed reduction.

First, propositions of law cannot be reduced to propositions of mo-

16. Consider G.E. Moore's objection to "naturalism." G.E. MOORE, PRINCIPIA ETHICA 18-21 (1903). Moore was a utilitarian; he believed the proposition that an action is right if it produces more good than any alternative action. If both "right" and "good" are understood as moral terms, then the proposition that Moore believed is an internal moral proposition. On the other hand, the further claim that the happiness of human beings constitutes the good is a bridge principle, relating a moral term ("good") to a nonmoral term ("happiness"). It is or aspires to be a bridge principle. What Moore referred to as the naturalistic fallacy was the belief that the *meaning* of "good" could be given by such nonmoral terms as "happiness," and that what I have called bridge principles could be used to define moral terms. If such terms could be defined that way, then all moral talk could be replaced by equivalent nonmoral talk; we could replace talk about the good with talk about happiness. Moore's "open question" argument was intended to show that morality is autonomous and not reducible to talk about some other thing. Moore argued that if "good" just meant "happiness," then it would make no sense to ask whether happiness was good; since it does make sense, one cannot be defined in terms of the other. This is the sort of argument that persuades those who are already persuaded. *Id.*

17. Naturally many examples are going to be harder than these to classify. Consider "Provocation is not a defense to murder." That appears to be an internal sentence linking the legal notion of provocation with the legal notion of a defense to murder. Yet it might well vary in truth value from jurisdiction to jurisdiction; there is nothing to stop a court or a legislature from making provocation a defense. Is this an internal sentence that varies in truth value? Or is it not an internal sentence? I think that the answer is that it is not internal, and that provocation is not a legal notion. This is a bridge statement, connecting the facts (or nonlegal norm) of provocation with the legal notion of a defense. The sentence in question makes a significant factual claim. Internal statements, like tautological statements, have universal application precisely because they have so little to say. They will typically have to do with relating and distinguishing legal concepts.

rality. The simple reduction of law to morality would yield a very simple-minded sort of natural law theory indeed. Not everything that ought to be the case is, or even ought to be, law. A more perceptive version of natural law theory is this: That what *is the law* and what *ought to be the law* are the same thing. But that claim is false as well—certainly it is false as a claim about the meaning of legal statements. There are laws that simply ought not to be laws; but even if there were not, the mere possibility that there might be is enough to defeat a meaning claim. Moreover, the claim does not propose a genuine reduction: the relevant moral propositions involve reference to the law essentially ("what ought to be the law"), so that there is no reduction of legal statements.

Second, legal statements cannot be reduced to statements descriptive of institutional facts. For example, it might be proposed that to say that some statement is the law is just to say that it has been uttered under the appropriate circumstances by the appropriate body—and perhaps the appropriate circumstances and the appropriate body are in part identified by the attitude of the governed toward them. But although such a procedure may be useful in identifying the bridge statements that govern the application of legal concepts in a particular jurisdiction, it says nothing whatever about the internal statements of the law. The truth of such internal statements does not depend on any facts about the production of law in a particular place at a particular time.

For example, it is true in every jurisdiction that whoever is justified in his actions is not criminally liable. That is just part of the machinery that societies bring into play when they implement a system of criminal law—it is part of what it means to have a criminal law system. It is up to the legislating body whether anything at all will justify, and therefore not only may what justifies in one society differ from what justifies in another, but there could well be societies in which nothing justifies at all. And it is a normative question, of course, what *ought* to justify (in any given jurisdiction). But to say that something is justified entails that there will be no criminal liability.

This has nothing to do, of course, with the words used in particular jurisdictions. It is possible that the word "justification" may be used in some jurisdiction to connote some other concept entirely, and we should not expect that concept to be involved in the same relations our notion of justification is involved in. In fact part of the benefit of identifying the internal part of the law is to enable us to identify the correlative elements of the law of different jurisdictions by focusing on the relations each element enters into rather than the words used. Such translation from one jurisdiction to another, someone is sure to point out, cannot depend entirely on structural similarities; there will have to be some starting point,

some concepts we can translate from one jurisdiction to another without reference to structural relations. But that problem is a universal one, and should not detain us here.

I conclude that there is an area of legal knowledge that is not reducible to institutional facts. Although societies are free to create whatever institutions they wish, not every such institution will count as a legal system. A legal system cannot maintain that someone who is justified is criminally liable, and if there is an institution that does so, it is not a legal system. (I would prefer not to get hung up on a particular example, because there naturally will be some controversy about what belongs to the internal realm. If you do not accept the point about justification and liability, consider an institution that found only people who are legally innocent, in our sense, to be criminally liable, in our sense. I will not speculate about whether such a system could be maintained, but it would not be a system of criminal law.)

The distinction between internal statements and bridge statements thus sets off a realm—the internal legal realm—that is the particular province of the legal scholar. The study of bridge statements, on the other hand, is the province the legal scholar shares with the political scientist, the economist, the philosopher, and others. The internal realm marks off a descriptive study whose truths should not vary from jurisdiction to jurisdiction. Wherever there is justification in the law, it relieves criminal liability; should there be something called justification in some jurisdiction that does not relieve liability, then it is not justification at all but something else. If we want to know what that jurisdiction does with justification, we must look for the appropriate concept, whatever it is called there. The same is not true with the particular grounds of justification, which might vary from jurisdiction to jurisdiction; "lesser evils" might provide a justification in one jurisdiction but not in others. That is a normative decision made on policy grounds or moral grounds, and it may be a correct decision and it may not.

Thus it is possible to distinguish three separate studies: the descriptive study of internal statements, which should not vary from jurisdiction to jurisdiction; the descriptive study of bridge statements, which will vary from jurisdiction to jurisdiction; and the normative study of bridge statements, the study of what the true bridge statements in a jurisdiction ought to be. If you believe that the acceptability of a bridge statement can be considered in isolation, then you may believe that the same bridge statements ought to be true in all jurisdictions. But if you believe (what is more likely to be true) that the desirability of any given bridge statement depends on the body of bridge statements it belongs to in the jurisdiction in which it occurs, then it may be that the normative study will

point to different bridge statements in different jurisdictions. I repeat that the descriptive and normative studies of the bridge statements would seem to presuppose some clarity about the formal relations among legal concepts.

C. *Some Observations*

(1) This description of formalism is merely a programme or a promissory note. The description does not attempt to identify particularly legal notions and distinguish them from the nonlegal, and clearly if it has not done that it has not performed the subsequent task of relating the legal notions to one another. There are two ways that we might set out to fulfill the promise. The first would be to try to formulate an entire logical system of legal expressions, working, so to speak, from the top down. Such a desire for systematization seems more Continental than Anglo-American. The other way would be to work at it piecemeal, marking off limited parts of the field to be worked at any one time. The choice between these two is a matter of taste, I think, and it may be worthwhile to have both sorts of endeavor going on at once. I must confess that the second of the two ways has more appeal for me, and I will undertake to provide an example of it in the second half of this Essay.

(2) Looking at things in terms of this distinction between internal and bridge statements throws some light on discussions in legal theory. For one thing, the formalism I have described is a genuine formalism, entailing as it does the autonomy of the law and the existence of a specifically legal realm of investigation. At the same time, it is not inconsistent with legal realism, since the existence of the bridge statements provides a connection between the law and other disciplines. For example, the observation a judge might make that it is legally unreasonable to require a marginal expenditure on care that exceeds the expected marginal gain in cost reduction is an example of a bridge statement, and it has the place in legal argument usually claimed for it by realists. On the other hand, internal statements connecting reasonableness and care with liability are required for any such argument to go through to an interesting legal conclusion.

(3) Formalism of this sort has nothing to say about the possibility of mechanical jurisprudence, a view expressed by Beccaria in the following paragraph:

> Judges in criminal cases cannot have the authority to interpret laws, and the reason, again, is that they are not legislators. . . .
>
> For every crime that comes before him, a judge is required to complete a perfect syllogism in which the major premise must be the general law; the minor, the action that conforms or

> does not conform to the law; and the conclusion, acquittal or punishment.[18]

The idea that laws made by human beings could cover every contingency, so that the only job for the judge is to pick out the applicable statute or precedent and apply it, deserves the abuse that it has gotten. Formalism does not entail this idea, but does not rule it out either. The bridge statements are rules of application created by courts or legislatures, and there is no reason to suppose that they cover every conceivable case, or even that they are not sometimes inconsistent. The internal statements, on the other hand, do not dictate the application of legal concepts.

(4) Traditional formalism is sometimes identified with natural law theory, which cannot be the case with the version I have described: if law is by definition a branch of morality, then it is not an autonomous discipline. Neither should formalism be identified with positivism, however; positivism has to do with the relationship between the descriptive study of bridge statements and the normative study of bridge statements. If positivism is true, then what is a bridge statement in a particular jurisdiction and what ought to be a bridge statement in that jurisdiction will be two different things.

II. THE FORMAL DISTINCTION BETWEEN JUSTIFICATION AND EXCUSE

Though this version of formalism makes claims that are very modest, its importance should not be underestimated. Discussions of legal theory may be confounded by failure to make the appropriate distinctions. I offer the example of the recent attempt to distinguish justification from excuse in criminal law. It is not so much that the distinction between internal and bridge statements can solve that problem as that this distinction can clarify just what the problem is.

An action violating the criminal law will not result in punishment if the action is justified or if the action is excused. Justified actions are sometimes said to be actions that we find desirable, like acting in self-defense; excused actions are actions that are not desirable but whose undesirability for one reason or another we cannot attribute to the actor, as in the case of insanity. There are other ways of marking the distinction between the two: for example, justification involves the objective rightness of the action, excuse the subjective state of the actor.

Kent Greenawalt has argued that various attempts to distinguish

18. CESARE BECCARIA, ON CRIMES AND PUNISHMENTS 14-15 (Henry Paolucci trans., Bobbs-Merrill 1963) (1784).

between the two notions must fail.[19] He has taken these attempts, including the two mentioned above, and has argued that each such attempt stumbles over the facts about what counts as justification and what counts as excuse. For example, the claim that justification looks to what is objective and excuse to what is subjective ignores the fact that in most jurisdictions justification depends upon the actor's perception of things, as well as the fact that the excuse of duress depends upon an objective standard: not what this actor is capable of, but what a person of ordinary firmness would be capable of. The claim that justification involves morally right actions while excuse involves morally permissible actions gets hung up on the difficulties of distinguishing between what is right and what is merely permissible.[20]

The problem is that many of the attempts Greenawalt discusses are apparently attempts to distinguish the bridge statements involving justification from the bridge statements involving excuse. The important point is that there is no necessary reason why the bridge statements—the principles of application of the terms—must differ and may not overlap. A particular jurisdiction may not distinguish the application of justification from the application of excuse, for example. Then of course there would be no clear line between the two sets of bridge statements. It does not follow that the concepts are not distinct or distinguishable, any more than the fact that "human being" and "featherless biped" may apply to all the same things implies that the two concepts are not distinct. Thus, when bridge statements are at issue, the enterprise is doomed to fail at the start.

Take the attempt to distinguish the two on a subjective/objective basis. It is a contingent matter, and up to the lawmaker, whether any particular jurisdiction will set an objective or a subjective standard for either justification or excuse. Jurisdictions will vary on their answer to this issue, and it is useless to suppose that the answer will provide a distinguishing mark between justification and excuse. Again, take the question whether justification applies only when the defendant chooses the lesser of the available evils. That cannot be a universal mark distinguishing justification from excuse since it is clearly up to the legislature whether that fact should justify at all or whether some lower standard is appropriate for justification.

---

19. Kent Greenawalt, *The Perplexing Borders of Justification and Excuse*, 84 COLUM. L. REV. 1897, 1899-1903 (1984).

20. For some reason that is not clear to me, Greenawalt seems to pull back from the initial intent of his article, which was apparently to show that it is impossible to draw a clear line between the two notions. Instead, he draws a more modest conclusion about what it is worthwhile for courts to be doing. *Id.* at 1927.

It is a different matter, I suppose, if the writers criticized by Greenawalt saw themselves involved in the normative criticism of bridge statements. If these writers are attempting to determine what the distinction between justification bridge statements and excuse bridge statements *ought* to be, then it is possible (though, as I have already observed, not necessary) that they might be looking for something that is the same in all jurisdictions. Nevertheless, even if they discover considerable overlap between justification and excuse in this regard, or find that there is no normative bridge characteristic that definitively distinguishes them, it does not follow that there is no *formal* distinction. And in fact the formal distinction is what the investigator must already have in hand when he begins either his descriptive or normative investigation of bridge principles. He must be able to *locate* in different jurisdictions those notions that we label justification and excuse.

The first appropriate search, therefore, is for the formal or internal distinction between justification and excuse. One of the distinctions discussed by Greenawalt seems to me to be of that sort. Fletcher and Robinson proposed that one's action is justified only if no one would be justified in resisting it, whereas one's action may be excused even though someone would be justified in resisting it.[21] Though this is a properly formal distinction—relating as it does the legality of one action with the legality of another—it is apparently simply false. For if it were true, then no one would be justified in resisting what reasonably appears to him to be unwarranted aggression simply because it does so appear to him. It must also in fact *be* unwarranted aggression. If the appearance of aggression were sufficient, as I think it is, then both parties might be justified,

---

21. *See* GEORGE P. FLETCHER, RETHINKING CRIMINAL LAW § 10.1.1, at 759-62 (1978); Paul H. Robinson, *A Theory of Justification: Societal Harm as a Prerequisite for Criminal Liability*, 23 UCLA L. REV. 266, 277-78 (1975). I feel sure that Fletcher and Robinson would reject the characterization of their work as formalistic. Nevertheless, some of the things Fletcher has to say about his own work and Robinson's suggest, at least partly, that this is the direction to be pursued. Fletcher said approvingly, for example:

> [Robinson] starts from a construction of the criminal law as a received set of principles that ought to be binding on our deliberations when conduct should be punished. The authority for these principles is not their instrumental value, but their grounding in a theory of just punishment *implicit in the patterns of liability that have accrued in the common law.*

George P. Fletcher, *The Right Deed for the Wrong Reason: A Reply to Mr. Robinson*, 23 UCLA L. REV. 293, 294 (1975) (emphasis added). Fletcher and Robinson clearly were not talking about something that happens to be true in different jurisdictions; and since they find it in the common law, neither can they argue that what they want is something that necessarily ought to be true in all jurisdictions, for that doesn't follow from the fact that it is in the common law. Insofar as they are clear about what they are looking for, it seems to me, they are looking for a formal distinction.

the first in attacking and the second in resisting, and according to the Fletcher-Robinson criterion that is conceptually impossible.

Another attempt depends on certain modal legal notions. We may say of a certain state of affairs that it is legally required to occur, that it is legally permitted to occur, or that it is legally forbidden.[22] And we might say that an action is justified if it is legally required, while an action is excused only if it is legally permitted, justification engendering our approval, excuse only our indulgence. Unfortunately, however, this also cannot be so. If a state of affairs is legally required, then its omission is legally prohibited. In the case of a justified action, to perform the legal action instead of the justified illegal action would be to omit the justified action, and if the justified action is required, that means that the legal action is legally forbidden. That cannot be the case. Whoever performs a justified action would have been entitled to perform the legal action instead. Both are permitted to him. Moreover, there may have been more than one action in a particular set of circumstances that would have been justified; but if a justified action is a required action, there can only be one justified action.

Let me propose, as a first approximation, a formal distinction between justification and excuse that involves a variation on the modal proposal and entails elements of the Fletcher-Robinson proposal as well. The proposal is this:

> *A*'s action is justified if and only if it occurs in circumstances (including the state of mind of the actor) such that anyone who performs it in those circumstances ought not, legally, to be punished; and the occurrence of the action is legally preferable, in the circumstances, to the occurrence of the legal alternative.

This distinction entails as much of the Fletcher-Robinson resistance thesis as is warranted: because the occurrence of the justified act is preferable to its nonoccurrence if the circumstances are as *A* believes them to be (*A*'s beliefs being one of the circumstances to be taken into account), the prevention of the justified act by someone who believes the circumstances to be as *A* believes them to be cannot also be justified. This is because the occurrence of the act and the prevention of the act cannot both be preferred in the same set of circumstances.

Moreover, to be preferable is different from being required: if an act is preferable in certain circumstances, its omission may still be permissible. Thus, even if breaking the law is justified it does not follow that

22. If we add that the state of affairs might be legally permitted not to occur, then we have a set of four predicates that fit into a classical logical structure known at least since Aristotle's time: the square of opposition.

obeying the law would be forbidden; it may well be permissible. And of course, unlike actions that are required, more than one action may be preferable to obeying the law.

In contrast:

> *A*'s action is excused if and only if it occurs in circumstances (including the state of mind of the actor) such that anyone who performs it in those circumstances ought not, legally, to be punished, and the occurrence of the action in those circumstances is *not* legally preferable to obeying the law.

In both cases—justification and excuse—the relevant circumstances may include the way the actor perceives the world around him, so that if he is mistaken about what is going on, that must be taken into account. In the case of involuntary intoxication it is not the circumstances as the actor believes them to be, but the actual circumstances that determine the outcome. Duress and related excuses are very similar to justifications in that it is the circumstances as the actor believes them to be that determines whether he ought to be punished; they differ in the manner of appraisal of the action itself.

What makes this distinction a formal one rather than a substantive one is that it defines both legal excuse and legal justification in terms of legal punishment and legal preferability rather than, for example, in terms of the proportion of harm to benefit created by the action. The decision as to what is legally preferable is one for the legislature and the courts: it may be legally preferable to create an excess of good over harm, and it may not be. But, assuming I have got the distinction right, it is not within the authority of the legislature or the courts to modify the distinction itself: if the relationship of excuse to justification is as I have described it, then it is part of the nature of the criminal law.

III. CONCLUSION

The important point is not the suggested distinction between justification and excuse, but rather the focus on the formal distinction rather than the applied distinction. In a jurisdiction in which no action is ever justified or excused, the range of application of the two notions will be exactly the same. It must be clear that sameness of application in that empty sense is irrelevant to the question of the distinction between justification and excuse. Or perhaps it is better to say that it is irrelevant to at least one important question about the distinction.

Is there any practical point to all this? One reason why Fletcher was interested in the relation between justification and excuse and the right to resist was that it showed that the distinction between justification

and excuse had a practical application. Is there a similar significance to be attributed to whatever turns out to be the correct formal distinction between the two? The formal distinction has profound practical significance, in a way I have already suggested: the very act of comparing justification in one jurisdiction with justification in another presupposes a good grasp of the logic and meaning of justification.