

Legal Formalism, Stage-Neutrality, and Comparative Justice

Prepared for presentation as a colloquium paper at the 2005 Meeting of the
American Philosophical Association, Central Division

Jeffrey Brand-Ballard, J.D., Ph.D.
523 Phillips Hall
Department of Philosophy
The George Washington University
Washington, DC 20052
(202) 994-6911
jbb@gwu.edu

I.

Several writers have argued recently that optimal rules of law authorize morally suboptimal decisions in certain cases.¹ Larry Alexander calls these “gap cases.”² Should judges in gap cases defer to legal rules or deviate from them? Philosophers known as “formalists” favor deference, “particularists” favor deviation.³

¹An optimal rule of law assigns legal relevance to all and only the correct case features. Writers who argue that optimal rules can authorize suboptimal decisions include Alan H. Goldman, *Practical Rules: When We Need Them and When We Don't* (Cambridge: Cambridge University Press, 2002) (hereinafter *PR*); Larry Alexander & Emily Sherwin, *The Rule of Rules* (Durham, NC: Duke University Press, 2001), ch. 4; Frederick Schauer, *Playing by the Rules: A Philosophical Examination of Rule-Based Decision-Making in Law and in Life* (Oxford: Clarendon Press, 1991), pp. 223-26; Larry Alexander, “Law and Exclusionary Reasons,” *Philosophical Topics* 18 (1990): 5-22; David Lyons, “Derivability, Defensibility, and the Justification of Judicial Decisions” in *Moral Aspects of Legal Theory* (Cambridge: Cambridge University Press, 1993), pp. 119-40, 128.

Alexander/Sherwin and Schauer hold that officials should defer to first-order legal rules. Goldman limits deference to higher-order rules including, for example, master rules such as H.L.A. Hart's ultimate rule of recognition. H.L.A. Hart, *The Concept of Law*, 2nd ed., (Oxford: Clarendon Press, 1994). Goldman endorses a metarule of judicial fidelity to positive law. *PR*, pp. 45, 46, 113, 130, 147.

²Larry Alexander coins the term in “The Gap,” *Harvard Journal of Law and Public Policy* 14 (1991): 695-701.

³Formalists include Joseph Raz, “Authority, Law, and Morality” in *Ethics in the Public Domain* (Oxford: Clarendon Press, 1994); *PR*; Alexander & Sherwin, *The Rule of Rules*; Schauer, *Playing by the Rules*; and Scott J. Shapiro, “Law, Morality, and the Guidance of Conduct,” *Legal Theory* 6 (2000): 127-70.

“Particularism” is the accepted term for those who favor deviation in at least some gap cases, used in Alexander and Sherwin, *The Rule of Rules*, p. 61; Schauer, *Playing by the Rules*, p. 97; *PR*, pp. 118, 187 n.50; and Gerald Postema, “Positivism, I Presume?,” *Harvard Journal of Law and Public Policy* 14 (1991): 797-822. It is a cousin of, though not to be confused with, what is called “particularism” in metaethics. See, e.g., Jonathan Dancy, *Moral Reasons* (Oxford: Blackwell, 1993).

I defend a position that partially supports and partially challenges formalism. I

distinguish between weaker and stronger formalist theses:

Weak formalism: for any type of gap case, judges ought to enforce the legal rules, rather than their own moral judgment, in at least some of the case tokens of that type.

Strong formalism: for any type of gap case, judges ought to enforce the legal rules, rather than their own moral judgment, in every case token of that type.

I support the weak thesis but not the strong. I argue that judges should deviate in a certain percentage of the tokens of a given gap-case type. Some of the arguments for the strong thesis look very persuasive at first, but I'll challenge two of these, advanced recently in Alan Goldman's book, *Practical Rules: When We Need Them and When We Don't*.⁴ Alan wants judges to adopt "as morally binding . . . a rule that they will follow settled law in their decisions, even when they disagree with the outcomes."⁵ Alan's arguments against my preferred method of selective deviation include an argument from comparative justice and an epistemic argument.⁶

I argue that Alan's justice argument is, interestingly, in tension with the standard argument for the weak thesis. Since the strong thesis subsumes the weak, an argument inconsistent with the weak thesis is therefore also inconsistent with, and can't be used to support, the strong thesis. This leaves the defense of strong formalism resting on the epistemic argument.⁷ Formalists have done little to substantiate the epistemic argument, however. Perhaps

⁴Cited supra, n.1.

⁵*PR*, p. 42. Goldman calls this a "fundamental rule of the legal system." *PR*, p. 147. He endorses it also on pp. 45, 46, 113, 130. He emphasizes, however, that this metarule does not require judges to follow first-order rules of law if they can cite a distinguishing feature of the case at bar, that is, a feature that a morally superior rule would treat as legally relevant. (More specifically, Goldman requires that such distinguishing be based on "values already recognized as legally relevant." *PR*, p. 137) However, such cases are, by definition, not gap cases. Gap cases only arise when the rule in question is already optimal (at least with respect to the case at bar). Goldman's position therefore entails strict prohibition on deviation in gap cases, though not strict prohibition on deviation in non-gap cases (i.e., those that present the need for a viable specification or elaboration of existing rules).

⁶*PR*, pp. 44-45. Goldman also presents a "social" or "coordination" argument which I address in a longer version of this paper.

⁷Notwithstanding the aforementioned coordination argument, which I also think can be challenged.

they've mistakenly assumed that the justice argument is available to them as a first line of defense.

II.

To appreciate what's at stake, consider the story of a judge whose philosophy reflects weak formalism, but not strong.

Once a month you visit the courtroom of your friend, Judge Holbrook. In July, Judge Holbrook hears the case of Jack, a teenager who has shoplifted jewelry from Wal-Mart. Jack has returned the jewelry. The statute classifies shoplifting as a Class A misdemeanor, minimum. Jack will lose his college scholarship if he's convicted of a Class A misdemeanor. To your surprise, Holbrook ignores the statute, convicting Jack of a lesser, Class B misdemeanor. Holbrook confides in you, privately, that the Class A conviction wouldn't have benefited anyone, including Wal-Mart (more than negligibly), and would have ruined Jack's future.⁸

You return to the courtroom in August to find Judge Holbrook hearing the case of Andy, another teenager who has shoplifted jewelry from Wal-Mart. Andy has returned the jewelry. A Class A conviction will cost him his scholarship. Having seen Jack's case, you're confident that Holbrook will convict Andy of a Class B. After all, Jack's and Andy's cases are tokens of a single legal case type.

To your astonishment, Judge Holbrook convicts Andy of a Class A misdemeanor. You privately question the Judge. Has he forgotten Jack's case? Was he privy to distinguishing facts of which you were unaware? Holbrook assures you that there were no distinguishing facts and that he remembers Jack's case clearly. In fact, he explains, in a sense he was hard on Andy *because* he was easy on Jack. He'd have liked to have given Andy the Class B as well, because

⁸I'm ignoring the likelihood that losing his scholarship might teach Jack a "valuable lesson" and thereby benefit him in some broader sense.

the Class A conviction benefited no one and greatly harmed Andy. But he hears cases of this type every month. He's confident that going easy on too many such defendants would, in the long run, cause systemic harm to the legal system that would outweigh the individual benefits of these lighter sentences. Other citizens would proceed to shoplift in non-optimal circumstances,⁹ mistakenly anticipating that Holbrook would go easy on them. It would also set a bad example for other judges, who might mistakenly go easy on such defendants. The aggregate harm to society of lenience for many such defendants outweighs the benefits.¹⁰ However, Holbrook explains, a judge can achieve the optimal effect by practicing lenience in a few such cases – neither too many nor too few.¹¹ Deviant lenience at the optimal level of density¹² produces a net benefit to society. Holbrook estimates that level as fifteen percent. From all the cases of the Jack/Andy case type, he randomly chooses fifteen percent of the defendants for lenience.

Judge Holbrook uses what game theorists call a *mixed strategy*.¹³ In my terms, he's a *holistic particularist*,¹⁴ albeit a candid and ham-fisted one.¹⁵ Holistic particularism invites

⁹By “non-optimal circumstances” I mean those in which the defendant believes lenience is justified but it actually is not.

¹⁰See *PR*, pp. 48.

¹¹See Donald Regan, *Utilitarianism and Cooperation* (Oxford: Oxford University Press, 1980), p. 62. Holbrook's strategy resembles, on an individual scale, the idea of a “small percentage of agents” breaking the rules in gap cases. *PR*, p. 59.

¹²Consider any region of space and time containing a legal system, such as Connecticut from 1990 to 1999. Compute the *deviation density* of that spatiotemporally delimited legal system as the fraction of cases in which deviation occurs, relative to the total number of cases heard in that system, and weighted, perhaps, by the average degree of deviation per case. David Lyons uses the term “density” similarly in *Forms and Limits of Utilitarianism* (Oxford: Clarendon Press, 1965), p. 72.

¹³See, e.g., R. Duncan Luce and Howard Raiffa, *Games and Decisions* (New York : Dover Publications, 1989), p. 70. Goldman considers mixed randomizing strategies. *PR*, p. 44. These strategies have effects similar to what Ronald Dworkin condemns, at the level of legislation, as “checkerboard solutions” to legislative controversies. *Law's Empire*, pp. 178-84, such as allowing abortions only to women born on an odd day of the month.

¹⁴Holbrook is a particularist because he sometimes deviates from justified legal rules in gap cases, based on the particular circumstances. But he is a “holistic” particularist because when he decides whether to deviate he factors in the systemic effects of deviation and his legal system's current overall level of deviation. Holistic particularism is a refinement of the view known in the literature as “rule-sensitive particularism.” Alexander and Sherwin, *The Rule of Rules*, p. 61; Schauer, *Playing by the Rules*, p. 97; *PR*, p. 118; Gerald Postema, *Bentham and the Common Law Tradition* (Oxford: Clarendon Press, 1986); and Postema, “Positivism, I Presume?” Goldman rejects rule-sensitive particularism. *PR*, p. 43.

¹⁵Holbrook's rule of thumb is extremely crude. In any given year, Holbrook may fail to optimize by following his

reactions ranging from reluctant approval to outrage.¹⁶ Alan sees “insuperable” problems with mixed strategies.¹⁷ If he’s right, then holistic particularism is indefensible.¹⁸

III.

Alan’s first problem with mixed strategies appeals to a cherished principle of common-sense morality, *comparative justice*.¹⁹ It’s often stated, in Aristotelian terms, as requiring us to “treat like cases alike.”²⁰ Alan describes it as the “Kantian constraint not to judge cases differently without finding a morally relevant difference between them.”²¹ Comparative justice also receives great attention and deference in Anglo-American jurisprudence.²² In jurisprudence,

rule. But that’s the nature of rules. In the long run, following rule of this type could have optimal results. Moreover, a more sophisticated judge wouldn’t wear his mixed strategy on his sleeve. He’d disguise his strategy under a veil of specious legal distinctions. Judges have many techniques available to help them conceal their violations of comparative justice. Compare this commentary by Henry Richardson:

Theoretically, an optimal intermediate solution may exist that would evict only a certain percentage of such widows – either on the basis of an arbitrary randomizing device or (more likely within the law) on the basis of a more finely specified principle; but Goldman, unlike some of the great jurisprudential thinkers of the past, does not trust the genius of the common law to evolve towards such solutions.

Henry S. Richardson, Review of Alan H. Goldman, *Practical Rules*, *Notre Dame Philosophical Reviews* (July 2002), <http://ndpr.icaap.org/>

¹⁶Over and above the outrage some will express at Holbrook’s initial contravention of the statute in Jack’s case.

¹⁷*PR*, p. 44.

¹⁸Interestingly, with respect to prudential reasoning Goldman supports a position similar to holistic particularism. *PR*, pp. 65-66, 70-72, 75-76, 96. He sees moral, epistemic, and coordinative problems in the interpersonal case that don’t arise in the intrapersonal case. *PR*, p. 73-74. This paper argues that these problems may not be fatal in the interpersonal case, either.

¹⁹Comparative justice goes by many names: “justice,” “fairness,” “equality,” “formal justice,” “formal equality,” et cetera. Goldman uses this terminology (*PR*, p. 167), but more often refers to comparative justice as the “Kantian constraint.” *PR*, pp. 3, 9, 22, 41, 79, 138-39, 140, 143-44, 159, 161, 167.

²⁰See, e.g., William K. Frankena, “The Concept of Social Justice,” in *Social Justice*, Richard Brandt, ed. (Englewood Cliffs, N.J.: Prentice-Hall, 1962), pp. 9-13; “Non-comparative Justice,” *Philosophical Review* 83 (1974): 297-338; S.I. Benn and R.S. Peters, *The Principles of Political Thought* (New York: Free Press, 1965), pp. 127-28; James I. MacAdam, “The Precepts of Justice,” *Mind* 77 (1968): 360-71; Michael Bayles, *Procedural Justice: Allocating to Individuals* (Dordrecht: Kluwer Academic, 1989).

²¹*PR*, p. 41.

²²Hart, *The Concept of Law*, 2nd ed., Penelope A. Bulloch and Joseph Raz, eds., (Oxford: Clarendon Press, 1994), p. 157 (noting the “peculiarly intimate connection” between justice and law); David Lyons, “On Formal Justice” and “Formal Justice and Judicial Precedent,” both in *Moral Aspects of Legal Theory* (Cambridge: Cambridge University Press, 1993), pp. 13-40, 102-118; Ronald Dworkin, *Law’s Empire* (Cambridge: Belknap Press, 1986), p. 183; Peter Westen, *Speaking of Equality: An Analysis of the Rhetorical Force of ‘Equality’ in Moral and Legal Discourse* (Princeton, N.J.: Princeton University Press, 1990).

the principle requires officials to treat identically all tokens of one case type.²³ Holbrook blatantly disregards this principle.

If comparative justice is valid then I can't defend weak formalism without strong. Weak formalism permits judges to deviate in some tokens of a gap case type, but not others. If comparative justice is valid, then deviation is permissible in either all of the tokens of a type, or none of them. Since I'm not defending the particularist claim that deviation is permissible in all tokens of a gap case, I must argue for qualifying or reinterpreting our commitment to comparative justice.

Comparative justice has *pro tanto* moral weight. *Ceteris paribus*, judges ought to obey it. I'll argue that all things aren't equal. There's a tension between comparative justice and the standard argument for weak formalism. To understand the tension we need a distinction between *stage-neutral* and *stage-relative* principles, one that refines the more familiar distinction between agent-neutral and agent-relative. A principle is agent-neutral if it prescribes the same ultimate ends to each agent, otherwise it's agent-relative.²⁴ A principle is stage-neutral²⁵ if it prescribes the same ultimate ends to each *temporal stage* of each agent, otherwise it's stage-relative.²⁶

²³Two tokens belong to the same case type if and only if there are no differences between them that would be treated as legally relevant by the optimal legal rules. On pain of trivializing the concept of "case type," we must understand this property as supervening on exclusively *local* properties of a case token. The fact that Jack's case was heard before Andy's case, for example, isn't a local property of either case, so it isn't part of the case type to which these tokens belong. Otherwise, Judge Holbrook's decisions don't violate comparative justice, as I've defined it. Some indirectly pertinent discussion appears in Shelly Kagan, "Causation, Liability, and Internalism," *Philosophy and Public Affairs* 15 (1986): 41-59.

²⁴Seminal sources include Samuel Scheffler, *The Rejection of Consequentialism*, 2nd ed. (New York: Oxford University Press, 1994); Thomas Nagel, *The View From Nowhere* (New York: Oxford University Press, 1986); Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), ch. 4. See also Stephen L. Darwall, "Agent-Centered Restrictions From the Inside Out," *Philosophical Studies* 50 (1986): 291-319; Christine M. Korsgaard, "The Reasons We Can Share: An Attack on the Distinction Between Agent-Relative and Agent-Neutral Values," *Social Philosophy and Policy* 10 (1993).

²⁵All agent-neutral principles are stage-neutral, so the term "stage-neutral" exists only for the sake of symmetry with "stage-relative."

²⁶Philosophers now recognize that the term "agent" in the definition of agent-relativity is ambiguous between an agent over the whole course of her life and an agent over some briefer timespan. Stage-relativity is a specification

There are both stage-relative and stage-neutral ways to understand comparative justice, the natural interpretation being stage-relative. So interpreted, the principle directs each judge, on each occasion, to compare the case before her to others she's heard. She's to rule in the instant case as she did in all other cases belonging to the same legal case type.²⁷

A stage-neutral interpretation of comparative justice, by contrast, requires judges to minimize the overall number of occasions on which a case is treated differently than others of its type. In certain situations, perhaps unlikely ones, this interpretation could require a judge *herself* to treat differently two cases of the same type, as a means to that stage-neutral end.²⁸

The defenders of comparative justice seem to intend the stage-relative interpretation.²⁹ The question now arises: can one consistently embrace both the standard argument for the weak formalist thesis and the stage-relative interpretation of comparative justice?

This seems unlikely, because the standard argument is patently stage-neutral. It basically runs as follows.³⁰ Deference to rules in gap cases leads judges to act wrongly on certain occasions, on some of which their individual judgment would lead them to act rightly. However, by obeying good rules, judges reach the right results more often than they would by simply exercising individual judgment in each case.³¹ This is because judges' individual judgment is costly to exercise, error-prone, lacks coordination with others, et cetera.³² Therefore, formalists

of agent-relativity that resolves that ambiguity. The point was made early in Richard Brook, "Agency and Morality," *Journal of Philosophy* 88 (1991): 190-212. See also, Robert H. Meyers, *Self-Governance and Cooperation* (Oxford: Oxford University Press, 1999); Christopher McMahon, "The Paradox of Deontology," *Philosophy and Public Affairs* 20 (1991): 350-77.

²⁷We might also understand the principle as requiring judges to compare the case to those she anticipates hearing.

²⁸These are "minimizing violations." Kasper Lippert-Rasmussen, "Moral Status and the Impermissibility of Minimizing Violations," *Philosophy and Public Affairs* 25 (1996): 333-350.

²⁹Goldman invites the stage-relative interpretation by repeatedly referring to CJ as the "Kantian constraint." See *supra*, n.18. He also states that we must use "deontological reasoning" when faced with "justified rules." *PR*, p. 55.

³⁰I ignore many nuances that do not bear on my argument.

³¹"In order to generate a rule, the cumulative effects of everyone's weighing moral reasons directly must be bad." *PR*, p. 48 (emphasis omitted)

³²See, e.g., *PR*, pp. 32-42; Alexander and Sherwin, *The Rule of Rules*, p. 54; Conrad Johnson, *Moral Legislation* (Cambridge: Cambridge University Press, 1991), ch. 3; Schauer, *Playing by the Rules*, ch. 5.

conclude, judges are rationally and morally obligated to defer. As Alan writes, “In applying [rules], we allow for less than the best outcomes in particular cases in order to achieve a cumulatively better state in the long run.”³³ This is the standard case for weak formalism. It’s a stage-neutral argument. It attributes to each stage the ultimate end of minimizing morally incorrect decisions by that stage and other stages.³⁴ This end contrasts with stage-relative ends such as *making morally correct decisions*, or *refraining from making morally incorrect decisions*. A stage-relative principle could, in theory, direct a judge to make a morally correct decision in one case, even if she thereby leads herself or her colleagues to make various morally incorrect decisions of equal import. Stage-relative principles instruct the agent to focus on her own agency at a particular time: what she does and intends, whom she uses as a means, not what she allows, foresees, or causes as an effect or aspect of her action.³⁵

IV.

How might one reconcile comparative justice with the argument for weak formalism? One might argue that comparative justice, even on a stage-neutral interpretation, will almost never permit judges to treat legally identical cases differently. Such violations are almost never required, in the real world, to prevent other such violations. Therefore, we can support the standard argument for weak formalism and a similarly stage-neutral interpretation of comparative justice to take us from weak formalism to strong, the critic concludes.

The critic correctly observes that violations of comparative justice are almost never required, in the real world, to minimize other such violations. They are, however, required to

³³PR, p. 64. See also p. 145 (“We need rules when the cumulative effects of direct reasoning that takes all normatively relevant factors into account are worse than the imperfect justice of the rules”); p. 81 (moral rules are “second-best strategies to *maximize* satisfaction of rights or other moral demands” [emphasis added]); and pp. 48, 147.

³⁴It makes no difference whether we understand the stage as minimizing incorrect decisions by all stages of the agent to which it belongs, or all stages of all agents.

³⁵Here I simply allude to some standard stage-relative distinctions from the deontological literature. See, e.g., F.M. Kamm, *Morality, Mortality*, vol. II (Oxford: Oxford University Press, 1996).

minimize moral wrongdoings by judges. That follows from Alan's original hypothesis: gap cases are those in which the judge does moral wrong by following a morally justified rule. Minimizing such wrongdoings requires judges to deviate³⁶ in some, though not necessarily all, gap cases of a given type, thereby violating comparative justice.

The critic might attempt to block this move by arguing that violations of comparative justice fall in a different, and lexically prior, category from other moral wrongs. The formalist argument, she might observe, concerns *moral* wrongdoing. I define comparative justice, by contrast, as concerning *legal* case types. Perhaps it's wrong to trade off moral wrongs against legal wrongs.

This solution, however, implies that comparative justice enjoys lexical priority over other values. Real legislators don't believe this. If they did, they could easily give judges greater lawmaking authority. The more lawmaking authority judges have, the easier it becomes for judges to *make it the case* that any two case tokens are legally identical or non-identical, leaving judges legally free to dispose of cases arbitrarily *without* violating comparative justice. A legislator who gave absolute priority to comparative justice over all other values would give judges just this power.

Real legislators don't give judges this power, thank heaven. The social costs of doing so would be unacceptable. This suggests that comparative justice does not, in fact, enjoy lexical priority over all other values. It's one value among many, as we might have suspected. *Ceteris paribus*, judges should treat identically cases that are legally identical. But all things aren't always equal. comparative justice isn't sufficiently important to justify granting judges broad

³⁶I use 'deviate' as a term of art meaning "to issue a decision in opposition to legal rules in a gap case."

lawmaking authority. At the other extreme, I suggest, it's not sufficiently important to justify the conclusion that judges are categorically wrong to deviate in gap cases.³⁷

I propose that supporters of the weak formalist thesis should abandon the stage-relative interpretation of comparative justice in favor of an indirect, stage-neutral interpretation. This allows them to harmonize their stage-neutral defense of the weak thesis with at least a qualified principle of comparative justice. A stage-neutral theory introduces the possibility that violations of comparative justice might be permissible in certain gap cases, as Judge Holbrook claims.³⁸

V.

So far I've argued that deviating in some fraction of gap cases isn't, in fact, immoral or irrational. My arguments support holistic particularism – a position that could conceivably vindicate Judge Holbrook – as a criterion of rightness, but they don't reach the viability of holistic particularism as a decision procedure.³⁹ I haven't established that legislators should explicitly *instruct* judges that deviation is ever permissible, nor that judges themselves should *believe* that it is. Thus we arrive at the next problem Alan sees for randomizing strategies, the “epistemic” problem. Each judge, he states, suffers “an inability . . . to identify the optimal pattern or to know how close to the threshold of damage to the legal system he might be in not enforcing the law.”⁴⁰ Because the judge can't know how close she is to the threshold,⁴¹ Alan concludes, it's irrational for her to deviate.

³⁷Goldman notes that strong rules can sacrifice moral consistency when they “prevent treating cases differently by blocking recognition of factors that differentiate” those cases. *PR*, p. 34. He notes that following rules can be inconsistent with integrity. *PR*, p. 77-78. But Goldman's own metarule favoring judicial fidelity is, by his own account, a strong rule. *PR*, p. 46. As such, I suggest, it threatens other values. That's the essence of my position.

³⁸Qualifying CJ doesn't mean abandoning it. In most cases, violations of CJ will not, in fact, promote the stage-neutral goal of minimizing wrongdoing. So CJ continues to have force across a wide range of cases, even on a stage-neutral interpretation. CJ has plenty of other important work to do in forbidding arbitrariness, partiality, and invidious bias. My suggestion concerns the exceptional cases in which CJ must give way to other values.

³⁹See R.E. Bales, “Act Utilitarianism: Account of Right-making Characteristics or Decision-making Procedure,” *American Philosophical Quarterly* 8 (1971): 257-65.

⁴⁰*PR*, pp. 44, 48, 53.

Alan's epistemic argument succeeds only if one of the following two assumptions holds:

1. Judges decide under conditions of uncertainty⁴²; OR
2. Judges decide under conditions of risk⁴³ and know that the probability of their legal system being at the threshold is relatively high.

Whereas, Alan's epistemic argument fails if

3. Judges decide under conditions of risk and know that the probability of their legal system being at the threshold is relatively low.

I submit that, in developed legal systems, the third assumption is true and the first two are false. Alan briefly considers this possibility, but his only responses are that, first, "the negative consequences of crossing [the threshold] are enormous" and, secondly, "a judge can assume that, in the absence of a rule, she or one of her colleagues will do so."⁴⁴

Alan's first response is true. Crossing the threshold could encourage lawlessness, with no compensating benefits. The truth of the second proposition, however, depends on the phrase "absence of a rule." If, for example, judges don't even understand that the threshold exists and blindly deviate in most gap cases, then someone will cross the threshold. In the real world, this isn't likely. Judges probably already understand that the threshold exists. If they don't already possess a rule for responsible deviation, I suggest one for them shortly.

First, I should note that such a rule is an epistemic possibility because the systemic harms of deviation have a special property, one that distinguishes them from the harms caused by more familiar patterns of collectively disadvantageous conduct. The systemic harms of deviation are *mind-dependent*. That is, deviation only generates systemic harm to the extent that at least some

⁴¹Define the *deviation density threshold* as the point of diminishing marginal returns: the level of deviation density such that any increase in density is expected to produce additional deviation in non-gap cases without additional benefits sufficient to compensate.

⁴²Meaning, they do not know even their probability of being at the threshold.

⁴³Meaning, they know their probability of being at the threshold, though they still do not know for sure whether they are at the threshold.

⁴⁴PR, p. 44-45.

people are aware of it. For example, Alan suggests that a pattern of rulings in which judges refuse to evict bankrupt widows who default on their mortgages would eventually make it more difficult for widows to obtain mortgages.⁴⁵ This is plausible, but only insofar as banks become aware of this pattern of rulings and conclude that it's too widespread to justify the risk of lending to widows.

Banks can, of course, research this sort of information. Aided by actuaries and lawyers, they may actually do so.⁴⁶ But there's no systemic reason why judges and legislators can't reproduce this research and learn what the banks learn. Courts may even be able to subpoena the banks' internal research to this end!

So here's a rule of thumb for judges, reflecting holistic particularism:

Before deviating in a case of type x , research the level of awareness, on the part of interested parties, of any existing pattern of deviation in x -type cases. In the absence of evidence suggesting such awareness, assume that deviating in the case at bar won't push one's legal system over the threshold. Proceed to follow a mixed strategy. Otherwise, defer to the rule.⁴⁷

VI.

It's instructive to compare the systemic harms of excessive deviation to more familiar collective harms, such as global warming, that are mind-*independent*. Burning coal was causing global warming long before anyone knew it, and I can't protect the planet by keeping my coal furnace a secret. Deviation, by contrast, causes no systemic harm unless people are aware of it.

This observation invites concerns that holistic particularism must be an esoteric⁴⁸ or "government house" theory.⁴⁹ Not so. My feedback argument actually relies upon the empirical

⁴⁵*PR*, p. 43.

⁴⁶Especially if the law permits them to rely upon such research when making lending decisions.

⁴⁷Alternatively, one could imagine a similar pattern emerging in a jurisprudentially heterogeneous legal system in which some judges pursue strong formalism while others tend more to particularism. The "genuine diversity in jurisprudential philosophy" that Goldman accurately observes (*PR*, p. 136) may achieve something like this result.

⁴⁸Henry Sidgwick, *The Methods of Ethics* (New York, Dover Publications, 1966), p. 490.

premise that judges' level of access to information won't lag too far behind that of legal actors who can be expected to adapt to patterns of judicial deviation. Most legal actors suffer from high levels of ignorance, in absolute terms, regarding patterns of judicial deviation. Holistic particularism simply treats this level of legal ignorance as a collective "resource" to be harnessed for social benefit.⁵⁰

The only systemic danger is that interested parties could become aware of an emerging pattern of judicial deviation *before* the judiciary itself does, and react to that pattern in a suboptimal way, for example, by denying loans to deserving widows. This would constitute a genuine collective action problem for judges. But the problem arises only if judges' individual decisions have the effect of accelerating aggregate deviation density⁵¹ at a rate that exceeds the system's capacity for self-equalizing feedback.⁵² It's an important question for empirical research under what circumstances this occurs. My point is that conceptual arguments, such as Alan invokes, can't do all the work.

Strong formalists need the comparative justice argument to support the conclusion that deviation is categorically wrong – that holistic particularism is indefensible even as a criterion of right. That argument conflicts with the standard argument for formalism itself, so strong formalists must retreat to the epistemic argument to support the conclusion that holistic particularism isn't viable as a decision procedure. The epistemic argument concludes that legislators should tell judges never to deviate and that judges should internalize these directions.

⁴⁹Bernard Williams, *Ethics and the Limits of Philosophy* (Cambridge: Harvard University Press, 1985), pp. 108-10.

⁵⁰Goldman himself notes the possibility of a "beneficial gap between public perception and actual practice." *PR*, p. 135, following on this point Meir Dan-Cohen, "Decision Rules and Conduct Rules: On Acoustic Separation in Criminal Law," *Harvard Law Review* 97 (1984): 625-77. A level of ignorance is probably inevitable in complex societies and is itself largely the product of invisible hand processes, rather than state coercion. Even those who are most fervently committed to democracy and the rule of law have rarely suggested that such "spontaneous" ignorance poses a threat to those values, however imprudent it may be for interested parties to remain uninformed.

⁵¹See definition of 'deviation density,' *supra*, n.11.

⁵²In other words, I suggest that judges *can* collectively monitor the cumulative effects of their collective decisions, in more or less the way Goldman believes to be impossible. *PR*, p. 74.

That argument may yet succeed, but it depends on empirical data that formalists haven't adequately marshaled.

VII.

Some readers might prefer to abandon weak formalism rather than qualify comparative justice as I've suggested. They would become stage-relative anti-formalists – radical particularists. On this view the judge strives neither to do nor to intend harm. She attempts, on each occasion, to issue the morally correct decision. She knows that, in doing so, she'll rule wrongly more often in the aggregate. But some will consider this an acceptable cost.

I don't think stage-relative anti-formalism reflects judicial practice or has much to recommend it. But I can't think of a philosophical objection that doesn't involve general objections to stage-relative theories. These would take me onto another subject. My view, however, is that the case others have made for weak formalism is sufficiently compelling to make worthwhile a reinterpretation of comparative justice.