Race and Criminal Justice Outcomes:

Triangulating Questions, Methods, and Measures

Samuel R. Sommers
Department of Psychology
Tufts University

Discussant Commentary
Symposium on the Past and Future of Empirical Sentencing Research
School of Criminal Justice, University at Albany
Baumer’s (this issue) review of empirical investigations of race and criminal sentencing nicely synthesizes the modal approach of previous research in this area and identifies future directions in the development of this line of inquiry. In this commentary, I elaborate on one of the recurring themes of Baumer’s piece, namely the importance of addressing race and criminal justice outcomes from multiple disciplinary and methodological perspectives. Such triangulation of research approaches is needed not only for the development of a clearer theoretical understanding of the processes that contribute to sentencing inequities, but also for efforts to persuade non-researchers who work in and evaluate the legal system that such disparities are real and merit intervention.

What do I mean by a triangulation of research approaches? Take, as one example, the present panel, in which two law school professors and a psychologist are commenting on the paper of a criminologist. In my case, I sit as a discussant in this conversation on race and sentencing despite having never actually conducted a single study on sentencing in my entire career. And I see this as a positive development. My research is on race and jury decision-making as well as race and jury selection, which means that I study some of the same processes that potentially contribute to sentencing inequities, just further upstream in the prosecution of cases. The very purpose of this symposium—to cast a more interdisciplinary eye on sentencing issues—is one that needs to be echoed in future efforts in this research area.
Indeed, again on a personal note, my experiences as a researcher and as an expert in court have repeatedly revealed the importance of addressing empirical legal questions from multiple points of view. Too often, individuals who are not trained researchers mistake a diversity of research approaches for confusion or capriciousness among researchers (to be fair, researchers themselves sometimes make this error as well). On the witness stand during cross-examination in capital trials, I have actually been confronted by the argument that the use of different methodologies to study the effects of race on legal outcomes—for example: archival analysis, juror interviews, and mock jury simulations—is a limitation of the literature. I’ve been asked to respond to questions along the lines of if researchers cannot even make up their minds regarding how to study these issues, how can you suggest that there is anything approaching a reliable set of empirical conclusions regarding the actual impact of race? If scientific consensus existed on the matter, this line of reasoning goes, then researchers would all be conducting similar-looking studies.

Lost in such arguments, of course, is the notion of convergent validity: the idea that when investigations making use of different methods, stimuli, and measures produce a similar set of findings, our confidence in the reliability of these findings increases rather than decreases. As students in my research methods courses hear me say week after week, multiple methods, stimuli, analyses, and measures are all good things. The rigorous researcher hopes to see such diversity when evaluating any line of inquiry.

Accordingly, I often begin my expert testimony (or my research talks to audiences outside of my discipline) by articulating exactly what this idea of convergent validity means in the domain of jury decision-making, and emphasizing the ability of one
methodology’s strengths to complement the limitations of another. Yes, the type of mock jury research that I conduct is inherently limited by external validity concerns. Mock jurors always remain aware that the judgments they render carry no real consequence, and this difference between mock and real jurors is a gap that can never be bridged entirely. Even after going to great pains to create a realistic and engaging trial simulation, the study remains just that: a simulation.

But the control and causal conclusions afforded to the researcher using experimental design are needed additions to a scientific body of knowledge that might otherwise rely solely on analysis of actual case outcomes and interviews with former jurors. And informative though these methods may be, there are limited by their inherently correlational nature: as fellow members of this panel can attest from personal experience, the conscientiously thorough researcher can identify and control for dozens (or even hundreds) of non-racial variables as potential confounds, and the skeptical critic is still able to come up with yet additional, unmeasured variables that supposedly cast doubt on the meaningfulness of the reported racial difference.

By the same token, the strengths of the archival analysis and juror interview methods speak directly to the limitations of the experiments I run. In particular, that these studies examine the actual decisions and outcomes that we researchers are ultimately interested in better understanding goes a long way towards addressing the criticisms of experimental work as too far removed from the reality of the courtroom to have any direct practical application. In short, there is no question that the research conclusion that race impacts the decisions of juries is much stronger for its foundation in
multiple study types than it would be living and dying with one methodology alone (Sommers & Ellsworth, 2003).

There are various ways to think about what this type of triangulation should look like in studying disparities in sentencing. As discussed above, one issue involves type of research method. Baumer (this issue, p. 25) suggests that “standard non-experimental regression approaches… have dominated the literature on race and sentencing.” In addition to future enhancement of the implementation of this particular method, he calls for a diversification of research approaches to studying sentencing outcome, though he also includes caveats warning of the practical obstacles that might impede such efforts.

I would frame the arguments in favor of this type of diversification even more forcefully, and would propose that some of the aforementioned obstacles are not as insurmountable as they may seem. For instance, Baumer writes (p. 27) that “while defendant race cannot in practice be experimentally manipulated… the perceptions of people who evaluate members of different racial groups can be randomized.” The impossibility of manipulating defendant race only applies, however, to researchers who remain within the confines of the modal approach of archival analysis via multiple regression. I would think it not controversial to suggest that mock juror experiments have contributed greatly to our understanding of the influence of race on juries (see Sommers, 2007; Sweeney & Haney, 1992) as well as a wide range of more general jury questions (see Devine, Clayton, Dunford, Seying, & Pryce, 2000); studies using experimental manipulation have also supplemented archival analysis of the impact of race on jury selection processes (see Baldus, Woodworth, Zuckerman, Weiner, & Brofitt,
2001; Sommers & Norton, 2008). I see no compelling reason why the same cannot hold true for experimental studies of the judgments involved in sentencing.

As just one example, Baumer quotes the conclusion of Paternoster and Brame (2008) that “it is not inconceivable to conduct an actual experiment in which prosecutors would be asked to decide whether they would seek a death sentence after reading a hypothetical case record of the homicide with the perceived race of the defendant and victim experimentally manipulated” (p. 28). Such a study is more than simply “not inconceivable”—it is precisely the type of investigation needed to better identify the root causes of sentencing disparities. Again, if experimental methods have been used to some success in answering other empirical legal questions regarding race, why can’t the same be accomplished when it comes to sentencing outcomes?

Of course, logistical obstacles remain. Baumer writes (p. 27) that “several of the sentencing scholars with whom I spoke expressed some skepticism about the viability of this type of research, owing largely to perceptions that prosecutors and judges might be reluctant to participate.” Such concerns are understandable. But persistent and creative researchers have been able to find ways to study real jurors, judges, attorneys, and even police officers. Certainly a reluctance to participate is encountered during, say, the effort to examine race and policing decisions, yet several psychologists recently have been able to secure access to police samples (e.g., Correll et al., 2007; Plant & Peruche, 2005). Moreover, the past few years have seen the establishment of a research-based Consortium for Police Leadership in Equity (www.cple.psych.ucla.edu). The CPLE is an interdisciplinary team of researchers and police officials cooperating on a range of research projects—on issues ranging from racial profiling to organizational equity to
immigration law enforcement—using police departments from more than a dozen sites across North America. While it would be foolish to suggest that sentencing research using actual attorneys and judges will be easy, precedent indicates that it is by no means an unattainable goal.

The advantages of experimental methods are not solely related to the ability to draw definitive conclusions regarding the causal relationship between variables (as important an advantage as that is). In addition, simulation studies afford researchers a better chance of identifying the motivations, attitudes, perceptual tendencies, and other psychological processes that underlie the outcome disparity often noted by archival analysis of real sentencing. To what extent to racially disparate sentences reflect conscious stereotypes versus more implicit forms of bias among attorneys, judges, or jurors? Do such discrepancies in sentencing reflect overly punitive decisions made for Black defendants, benefit-of-the-doubt leniency granted to White defendants, or some combination thereof? Questions like these—that unpack the main effect differences identified by archival analysis—are particularly amenable to controlled investigation using experimental methodology.

Beyond new research methods, another, even easier means of diversifying the empirical literature is simply pursuing new analytic strategies and techniques. In his review, Baumer writes (p. 21), “There are several possible meaningful mediators through which defendant race may affect sentencing outcomes, including pre-trial detention and bail amounts, mode of adjudication, and the quality of legal representation. Too often, these factors are considered merely as controls that must be included to yield a meaningful estimate of the direct effect of defendant race. Greater attention to the
mechanisms through which defendant race may matter in indirect ways is vital if we are to fully grasp its importance in shaping sentencing outcomes.”

This is a persuasive argument. At the same time, this glaring empirical lacuna seems easily bridged. If researchers adhering to the “modal method” of examining sentencing already assess quantitatively open-ended factors like these, then the transition from treating these variables as mere controls to examining them as predictors that interact with defendant race is a simple statistical matter. That is, if the data already exist, there is no reason to settle for a call-to-arms—these analyses can be conducted right now, unlike those future research directions requiring additional data collection. Such reanalyses may also identify promising mediational pathways that can later be examined in an even more direct, a priori manner via the type of experimental research discussed above.

And, of course, a diversified, triangulated approach to studying criminal justice outcomes would also devote additional attention to policies and processes occurring earlier “upstream” from final sentencing, as Baumer details in his concluding paragraphs. From the creation of laws and mandatory penalties to police enforcement priority decisions to prosecutorial discretion to jury composition and decision-making, there are many links in the chain of justice that can contribute to the sort of disparities at the heart of the matters considered by this panel. Many of these links are already the subjects of extensive bodies of research, yet crosstalk between those of us who study these different issues remains less consistent than it should be. Actual collaborative endeavors across research programs and disciplines may be even less common, yet even more important for future examination of our justice system.
There are both theoretical and practical future questions to be posed regarding race-related disparities in criminal sentencing. Much of my commentary has focused rather explicitly on the former—the effort to develop a more complete conceptualization of the extent to which race predicts sentencing outcomes and the processes by which such influence occurs. Baumer (p. 24) classifies the extant research agenda into four specific objectives: “(a) detecting racial disparities; (b) detecting racial discrimination, (c) evaluating whether a given policy intervention has modified observed racial disparities or discriminatory outcomes, and (d) assessing how race influences legal decision makers and/or legal decisions.”

But let us not forget that inherent to the study of sentencing disparities is the potential to lay the groundwork for real-world efforts to ameliorate the problem, whether in the form of amicus briefs, expert testimony, interventions, or procedural change. No, not all researchers in this area have (or must have) such goals. But each investigation contributes something new to our scientific body of knowledge—a body upon which other researchers and practitioners base attempts at calling attention to and remedying disparity. Therefore, triangulating the investigation of race and criminal justice outcomes serves more than just the theoretical or conceptual motivations highlighted by Baumer and emphasized in the present commentary. This objective is also a necessary precursor to developing more effective practical efforts to address sentencing disparities, whether generated by researchers or by other individuals based on this research.

That said, a more diversified body of empirical data is not, on its own, sufficient to ameliorate disparities in sentencing. As one example, my experience suggests that
convincing a trial judge to take the death penalty off the table in a particular case because research has identified general race-based disparities in the application of capital punishment is an uphill, if not fruitless climb. Even when supported by diverse empirical evidence of systematic bias, such a motion will almost inevitably fail because attorneys and experts cannot prove conclusively that the identified pattern of general discrimination in the scientific literature will definitively emerge in the case at hand. So a triangulated research literature is certainly no silver bullet that guarantees equitable legal outcomes.

But efforts to inform courts and policymakers about scientific research—or to effect change through any other means—have almost no hope at all when couched in a solitary methodology or perspective. Having all your research eggs in one basket is not a recipe for successful advancement of practical objectives because any one research method has its inherent limitations. Adhering rigidly to one modal strategy rarely makes for particularly good science, either. Ultimately, I agree wholeheartedly with Baumer that the most important message to be drawn from his review is the one he ends with: “the next generation of scholars interested in studying race and sentencing [must] engage deeply into the pertinent literatures related to the topic in sociology, psychology, law, economics, and other areas” (p. 35). Adopting and adapting these colleagues’ methodological tactics would be a most welcome development as well.
References


Baumer, E. P. (this issue). Reassessing and redirecting research on race and sentencing.


