REGULATING THE PRODUCTION OF KNOWLEDGE: RESEARCH RISK–BENEFIT ANALYSIS AND THE HETEROGENEITY PROBLEM

MICHELLE N. MEYER*

TABLE OF CONTENTS

* J.D., Harvard Law School; Ph.D., University of Virginia. Academic Fellow, Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School. For helpful conversations and comments on earlier versions of this Article, I thank Christopher Chabris, Glenn Cohen, Holly Fernandez Lynch, Michael Frakes, Janet Halley, Christine Jolls, Ben Roin, Jeff Skopek, Patrick Taylor, Melissa Wasserman, Alan Wertheimer, and participants in Harvard Law School's Health Law Policy Workshop. I am also grateful to the American Society of Law, Medicine & Ethics (ASLME) and the Saint Louis University Center for Health Law Studies, co-sponsors of the 2011 Health Law Scholars Workshop where this paper was selected for presentation, and to the Nominating Committee, readers Jesse Goldner, Efthimios Parasidis, Frank Pasquale, and Sidney Watson, and the other workshop participants. The usual caveats apply.

> The perception that agencies are out of control arises from the fact that in being called on to make fundamental value judgments they have moved outside their accustomed sphere of activity, outside their expertise, and outside the established system of controls. This perturbation of the regulatory process will not be corrected until the regulatory agencies are relieved of the necessity of making judgments they are not equipped to make.
>
> —Richard M. Cooper, Food and Drug Administration (FDA) Chief Counsel, 1978[1]

INTRODUCTION

Scholars and lawmakers expend considerable effort determining optimal incentives for innovation.[2] They expend similar effort ensuring that socially useful knowledge, once produced, is widely[3] and accurately[4] disseminated and implemented.[5] Yet, if knowledge-producing activities themselves are suboptimally regulated, neither upstream incentives to engage in them nor downstream mechanisms to disseminate their fruits will fully achieve their desired effects. And so it is both curious and problematic that the optimal regulation of knowledge-producing activities themselves is almost entirely neglected in this literature.

This Article critically examines the regulation of those knowledge-producing activities with the greatest potential to affect human welfare: research involving human beings, or "human subjects research" (HSR).[6] A single, neglected regulatory framework adopted by more than one dozen federal agencies[7] governs the production of the vast majority of our most important knowledge—from drug trials, to quality improvement

---

1. Richard M. Cooper, *The Role of Regulatory Agencies in Risk-Benefit Decision-Making,* 33 FOOD DRUG COSM. L.J. 755, 772 (1978).

2. In legal scholarship, innovation policy is virtually synonymous with intellectual property law in general, and with patents in particular. Other incentives include market exclusivity, trade secrets, prizes, research grants, and subsidized education.

3. For example, First Amendment law and free speech norms, education law, mandatory disclosure rules, data sharing rules and open source norms, patent and copyright limits and reversion to the public domain, patent disclosure, compulsory licenses, and fair use.

4. For example, prohibitions on false or misleading information, regulation of labeling, and libel and defamation law.

5. For example, direct funding of translational science and subsidized education in translation-relevant sciences.

6. Although an individual about whom research is conducted is traditionally called a "subject," I follow more recent usage and refer to research "participants." I revert to "subject" only when quoting other sources, when use of "participant" would be ambiguous, or when referring to "human subject research" (HSR), a locution that has not evolved to reflect the change from "subject" to "participant."

7. *See infra* note 26.

studies designed to reduce medical errors, to policy experiments that test the effectiveness and efficiency of governmental programs and regulations, to studies of the causes and effects of cognitive biases and implicit bias. Although this Article refers to U.S. regulations for convenience, its argument is equally applicable to the research governance of most other industrialized (and, increasingly, developing) countries.[8] As a result, the heterogeneity problem is equally applicable to these governance systems.[9]

The Article focuses, furthermore, on the primary actors in the regulation of HSR—licensing committees called Institutional Review Boards (IRBs), which, pursuant to federal statutes and regulations, review and must approve each study before it may proceed. Although the regulation of HSR is largely overlooked by scholars of innovation policy, this Article is hardly the first to critique IRBs. For decades, critics of IRBs have tended to fall within one of two broad camps.[10] One camp, comprised chiefly of bioethicists[11] who, in this context, appeal to deontological norms such as justice and anti-exploitation, charges IRBs with *underregulating* research. Whether the culprit in their eyes is institutional capture, conflicts of interest, or insufficient expertise, training, and material resources, these critics argue that IRBs are prone to Type I errors, which allow unreasonably risky research to proceed.[12] By contrast, a much smaller camp, comprised chiefly of scholars of regulation, governance, and bureaucracy,[13] appeals largely to economic efficiency in arguing that IRBs *overregulate* by rejecting, altering, and delaying reasonable research. Critics say these Type II errors impose administrative, opportunity, and academic freedom[14] costs to

8. *See infra* text accompanying notes 169–71.

9. *See* H.E.M. van Luijn et al., *The Evaluation of the Risks and Benefits of Phase II Cancer Clinical Trials by Institutional Review Board (IRB) Members: A Case Study*, 32 J. INST. MED. ETHICS 170, 174 (2006) ("[T]he structure, objectives, and procedures of IRBs are similar, regardless of whether they are American or European."). Not surprisingly, then, numerous other countries have reported similar problems with their research ethics review systems, including costs, delays, inconsistency, lack of transparency, and disproportionate regulation relative to risk.

10. *See* Ezekiel J. Emanuel & Jerry Menikoff, *Reforming the Regulations Governing Research with Human Subjects*, 365 NEW ENG. J. MED. 1145 (2011).

11. I include here those who approach HSR from the perspective of law and bioethics.

12. *See, e.g.*, Carl H. Coleman, *Rationalizing Risk Assessment in Human Subject Research*, 46 ARIZ. L. REV. 1, 2 (2004) (noting cases where otherwise healthy individuals died during HSR, which "turned out to be symptoms of deep and pervasive problems" and an "unprecendented crisis" in "our system of protecting human subjects").

13. These scholars are joined by some First Amendment scholars, *see infra* note 14, and by legions of disgruntled researchers.

14. *See, e.g.*, Philip Hamburger, *Getting Permission*, 101 NW. U. L. REV. 405 (2007). *But see* Richard A. Epstein, *Defanging IRBs: Replacing Coercion with Information*, 101 NW. U. L. REV. 735 (2007) (IRBs are constitutional); John A. Robertson, *The Social Scientist's Right to Research and*

researchers and institutions, as well as health and other costs to society from delayed, blocked, or foregone knowledge production, all of which far outweigh any benefits to participants of IRB review.[15]

IRBs are legally required to approve only those studies whose "[r]isks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result."[16] This risk–benefit analysis is therefore critical to the goals of both camps of critics—that is, to preventing reasonable research from being blocked, altered, or delayed, and to preventing unreasonable research from proceeding.[17] Yet "determining risk–benefit ratios is one of the most important but least developed areas of determining the ethics of research trials."[18] To the extent that both camps indirectly address IRBs' risk–benefit standard, they assume the very premise that this Article argues is unsound. Charges of both Type I and Type II errors, as well as the most popular and seemingly promising proposals to correct these errors, assume that IRBs, regulators, and their critics are capable of correctly determining the single risk–benefit profile of each study for each and every potential participant. This Article argues that what it calls the "heterogeneity problem" renders IRBs intrinsically incapable of meaningfully performing the risk–benefit analysis the regulations demand of them.

The heterogeneity problem has two facets. The first is informational. Both the probability and magnitude of risks and expected benefits—as well as the "reasonableness" of assuming a particular bundle of risks in pursuit

---

the IRB System, *in* ETHICAL ISSUES IN SOCIAL SCIENCE RESEARCH 356 (Tom L. Beauchamp et al. eds., 1982) (same); James Weinstein, *Institutional Review Boards and the Constitution*, 101 NW. U. L. REV. 493 (2007) (same).

15. *See, e.g.*, Scott Burris, *Regulatory Innovation in the Governance of Human Subjects Research: A Cautionary Tale and Some Modest Proposals*, 2 REG. & GOVERNANCE 65, 67–68 (2008); Scott Burris & Jen Welsh, *Regulatory Paradox: A Review of Enforcement Letters Issued by the Office for Human Research Protection*, 101 NW. U. L. REV. 643, 645 (2007); David A. Hyman, *Institutional Review Boards: Is This the Least Worst We Can Do?*, 101 NW. U. L. REV. 749, 753, 756 (2007); Todd J. Zywicki, *Institutional Review Boards as Academic Bureaucracies: An Economic and Experiential Analysis*, 101 NW. U. L. REV. 861, 866, 875, 883 (2007).

16. 45 C.F.R. § 46.111(a)(2) (2012).

17. Not surprisingly, an IRB's risk–benefit analysis correlates strongly with its ultimate decision as to the acceptability of the research. van Luijn et al., *supra* note 9, at 172.

18. Manish Agrawal & Ezekiel J. Emanuel, *Ethics of Phase 1 Oncology Studies: Reexamining the Arguments and Data*, 290 J. AM. MED. ASS'N 1075, 1077 (2003); *see also* Coleman, *supra* note 12, at 4 & n.16; C. Lenk et al., *Non-Therapeutic Research with Minors: How Do Chairpersons of German Research Committees Decide?*, 30 J. INST. MED. ETHICS 85, 85–86 (2004); Annette Rid & David Wendler, *Risk-Benefit Assessment in Medical Research—Critical Review and Open Questions*, 9 L. PROBABILITY & RISK 151, 174 (2010); Douglas K. Martin et al., *The Incommensurability of Research Risks and Benefits: Practical Help for Research Ethics Committees*, IRB: ETHICS & HUM. RES., Mar.–Apr. 1995, at 8.

of a particular bundle of potential benefits for oneself or others—depend significantly (though not exclusively) on the preferences and other personal circumstances of individual prospective participants. Because IRBs assess risks and benefits before individual prospective participants are even identified, they lack access to these critical inputs. IRBs thus face a classic "central planner's problem": they are charged with making decisions about the acceptability of research based largely on participants' welfare, yet much of the information necessary to meaningfully predict the extent to which research participation would further or set back participants' interests is local information that resides with prospective participants, not IRBs.

Moreover, even if IRBs could solve their information problem, they face a second problem: aggregation. Because of prospective participant-will heterogeneity, a study that imposes a "low" risk on one participant will likely impose a "high" risk on another—and an expected benefit on still another. Finally, even if all prospective participants could expect the same costs and benefits from participating in a study, they are very likely to differ in their willingness to assume those risks in pursuit of those benefits. Yet IRBs must assign a single risk–benefit profile to each study, and then determine, for all prospective participants, whether that risk–benefit profile is "reasonable."

This Article proceeds in four parts. Part I offers a brief overview of the procedural and substantive rules that govern HSR and of the surprisingly broad range of actors and activities to which they apply. Part II shows how participant heterogeneity renders impossible the well-intentioned attempts of IRBs to determine, for each and every potential participant, a study's risks, its expected benefits, and the "reasonableness" of the former relative to the latter. This Part draws on empirical research in several fields that finds considerable individual differences in susceptibility to a variety of research-related harms and in benefits and risk–benefit tradeoff preferences.

Although regulators, no less than academic commentators, generally fail to acknowledge participant heterogeneity, the IRB system has developed strategies for assessing risks and expected benefits that implicitly respond to participant heterogeneity, and take very different forms in the risk and benefit contexts. Part III articulates these strategies and argues that their net effect is significant IRB risk aversion relative to the preferences of many—and in the case of some studies, likely most—prospective participants. This Part then describes some of the costs of this risk aversion.

Part IV considers several popular, seemingly promising proposed reforms of IRBs, many of which are designed to address perceived "errors" in IRB risk–benefit analysis. It argues that none would significantly

mitigate, much less solve, the heterogeneity problem, and that some would exacerbate it. The Article concludes by sketching the broad policy choices we face in light of the intractability of the heterogeneity problem.

I. BACKGROUND

A. Statutory Basis for IRB Review

Title II of the National Research Act of 1974,[19] Protection of Human Subjects of Biomedical and Behavioral Research, covers any entity applying for a grant or contract to conduct research involving humans under the Public Health Services Act. It requires those entities to provide "assurances satisfactory to the Secretary" of the (then) Department of Health, Education and Welfare that the entity has established an IRB to review that research "in order to protect the rights of the human subjects."[20] The Act directed the Secretary to promulgate within 240 days regulations pertaining to IRBs and assurances,[21] and established the Office for Protection from Research Risks (OPRR), an agency within the National Institutes of Health, to oversee assurances and IRBs. In 2000, OPRR was renamed the Office for Human Research Protections (OHRP) and relocated to a more prominent, less capture-prone position within the Office of the Secretary of the Department of Health and Human Services (HHS).

The Act also established a powerful ad hoc commission, the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (Commission), to which Congress delegated most substantive policy questions regarding IRB review.[22] Among the items the Act directed the Commission to consider was the "role of assessment of risk–benefit criteria in the determination of the appropriateness of research involving human subjects."[23] Between 1975 and 1978, the Commission published a series of reports[24] that, with only modest changes, formed the

19. Pub. L. 93–348, 88 Stat. 342 (1974).

20. *Id.* § 212(a), 88 Stat. at 352–53.

21. *Id.* § 212(b), 88 Stat. at 353.

22. *Id.* § 201(a)–(b)(1), 88 Stat. at 348. For more on Congress's delegation to the Commission, see Michelle N. Meyer, Research Contracts: Towards a Paternalistic Market in Research Risks and Benefits (Aug. 29, 2012) (unpublished manuscript) (on file with author) (detailing Congress's delegation to the Commission).

23. §§ 202(a)(1)(B)(i)–(v), 202(a)(1)(C), 88 Stat. at 349; *see also* Meyer, *supra* note 22 (arguing that regulations requiring prospective third-party risk–benefit analysis was not responsive to the abuses in HSR that Congress had identified).

24. The most well-known of these is NAT'L COMM'N FOR THE PROT. OF HUM. SUBJECTS OF BIOMED. & BEHAVIORAL RES., THE PROTECTION OF HUMAN SUBJECTS OF

basis for HHS's 1981 regulations.[25] In 1991, in order to achieve a consistent federal policy on HSR, virtually every federal department and agency that conducts or funds HSR adopted HHS's regulations.[26] Officially entitled *The Federal Policy for the Protection of Human Subjects*, the regulations have, since their widespread adoption, been better known as the "Common Rule." At the same time, the FDA amended its regulations to conform as closely as possible to the Common Rule, commensurate with its enabling statute.[27]

RESEARCH: THE BELMONT REPORT (1978) [hereinafter BELMONT REPORT], but the Common Rule is essentially a codification of NAT'L COMM'N FOR THE PROT. OF HUM. SUBJECTS OF BIOMED. & BEHAVIORAL RES., INSTITUTIONAL REVIEW BOARDS: REPORT AND RECOMMENDATIONS (1978) [hereinafter NAT'L COMM'N, IRBS].

25. The Department of Health, Education and Welfare (HEW) promulgated the first federal regulations for the protection of subjects on May 30, 1974, just weeks before Congress passed the National Research Act, and codified them at 45 C.F.R. pt. 46 (1975). *See* Protection of Human Subjects, 39 Fed. Reg. 18,913 (May 30, 1974). In 1981, it revised them in light of the Commission's work. *See* 45 C.F.R. pt. 46 (1981).

26. *See* Federal Policy for the Protection of Human Subjects, 56 Fed. Reg. 28,003 (June 18, 1991), codified at 7 C.F.R. pt. 1c (2013) (Department of Agriculture); 10 C.F.R. pt. 745 (2013) (Department of Energy); 14 C.F.R. pt. 1230 (2013) (National Aeronautics and Space Administration); 15 C.F.R. pt. 27 (2013) (Department of Commerce); 16 C.F.R. pt. 1028 (2012) (Consumer Product Safety Commission); 22 C.F.R. pt. 225 (2012) (Agency for International Development); 24 C.F.R. pt. 60 (2012) (Department of Housing and Urban Development); 28 C.F.R. pt. 46 (2012) (Department of Justice); 32 CFR pt. 219 (2012) (Department of Defense); 34 CFR pt. 97 (2012) (Department of Education); 38 C.F.R. pt. 16 (2012) (Department of Veterans Affairs); 40 C.F.R. pt. 26 (2011) (Environmental Protection Agency); 45 C.F.R. pt. 46(A) (2011) (Department of Health and Human Services (HHS)); 45 C.F.R. pt. 690 (2011) (National Science Foundation); 49 C.F.R. pt. 11 (2011) (Department of Transportation); *see also* NAT'L BIOETHICS ADVISORY COMM'N, 2 ETHICAL AND POLICY ISSUES IN RESEARCH INVOLVING HUMAN PARTICIPANTS J–23 (2001) [hereinafter NBAC] (the Office of Science and Technology Policy has adopted the Common Rule but has not codified it in regulations); Social Security Independence and Program Improvements Act of 1994, Pub. L. No. 103–296, § 106(b), 108 Stat. 1464, 1474–76 (1994) (providing that the Common Rule continues to apply to the Social Security Administration after its 1995 secession from HHS); Exec. Order No. 12,333, 46 Fed. Reg. 59,941, 59,941 (1981) (making the Common Rule applicable to the Central Intelligence Agency). This Article cites the HHS regulations for simplicity. Other federal agencies and departments have adopted their own HSR regulations. *See* 34 C.F.R. § 356.3(c)(1) (2012) (Office of Special Education and Rehabilitation Services); 28 C.F.R. § 512.10 (Bureau of Prisons). A few states have statutes or regulations that govern HSR, either in general, see, e.g., CAL. HEALTH & SAFETY CODE §§ 24170–24181 (West 2012); N.Y. PUB. HEALTH LAW §§ 2440–2446. (McKinney 2012); Va. Code Ann. §§ 32.1–162.16. (2011); FLA. STAT. ANN. § 381.85 (West 2007) (repealed 2010), or in specific research contexts, see, e.g., CAL. HEALTH & SAFETY CODE §§ 24185–24189 (West 2012) (cloning research); OR. REV. STAT. §§ 431.805–431.815. (2011) (genetic research). Following convention, when referring to the Common Rule, this Article will henceforth cite to the HHS regulations.

27. *See* 21 C.F.R. pts. 50, 56 (2012).

B. Regulations Governing IRB Review

IRB review is designed to protect research participants, and IRBs approve, disapprove, or require changes to each study accordingly.[28] Before researchers recruit a single participant, IRBs review their recruitment plans, the detailed information disclosures that form the basis of participants' voluntary, informed consent, and the protocol itself. They ensure that these materials fully, accurately, and in "understandable" language disclose to prospective participants, *inter alia*, "any reasonably foreseeable risks or discomforts to [them]" and "any benefits to [them] or to others which may reasonably be expected from the research."[29] They then consider these risks and expected benefits themselves, and approve only those studies whose "[r]isks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result."[30]

Although many associate IRBs with biomedical research, in the United States, much industry and almost all academic HSR is subject to IRB review, either directly, through federal statute and regulations, or indirectly, through contract. Thus, IRBs license everything from Phase I trials of investigational new drugs to quality improvement activities and experimental economics and philosophy, to sociology surveys, oral history, and the studies that form the basis of the burgeoning empirical legal studies movement. Suboptimal regulation of HSR by IRBs thus has a substantial impact on knowledge production and participant welfare.

1. Covered Actors

By their terms, both the National Research Act and the Common Rule require IRB approval only of HSR conducted, funded, or otherwise subject to regulation by any Common Rule agency or department. In practice, however, a web of contractual relationships ensures that most HSR, including virtually all HSR conducted by academics and their students, is subject to IRB review[31] regardless of the source of funding.

The Act requires each institution engaged in federally funded HSR (for example, a university or academic medical center) to provide assurance that it will adhere to the regulations. The regulations implement this directive

28. 45 C.F.R §§ 46.109(a), 46.116(a) (2012).

29. *Id.* § 46.116.

30. *Id.* § 46.111(a)(2); *see also* 21 C.F.R. § 56.111(a)(2) (2012) (same IRB risk–benefit requirement applicable to HSR subject to Food and Drug Administration (FDA) jurisdiction).

31. Whether every researcher whose work is *subject* to IRB review in fact submits it to an IRB is a separate matter.

by requiring that each institution file a standard form contract between the institution and OHRP called a Federal Wide Assurance (FWA). The standard FWA invites the institution to apply the regulations to *all* HSR in which the institution is "engaged," regardless of whether that research receives federal funding or not.[32] This generally has been interpreted to extend to all HSR conducted by any of the institution's faculty or students. Historically, between 74% and 90% of institutions have agreed to this condition.[33] Since someone at virtually every academic institution receives federal funding for research, and since the vast majority of institutions agree to apply the regulations throughout their campuses, the overwhelming majority of HSR conducted in academic settings is subject to IRB review. Federal regulators are currently considering proposed reforms of the Common Rule under which federal funding of any investigator's research would be conditioned upon her institution extending IRB review to all HSR in which it is engaged.[34]

Even of those few institutions that do not contract with OHRP to extend IRB review, "many" nevertheless have adopted a policy under which they extend IRB review to all faculty HSR, student HSR, or both.[35] Similarly, many journals require that research submitted for publication be approved by an IRB.[36] Thus, if a researcher is not subject to IRB review directly through a federal grant or contract, she will likely be subject to it indirectly,

---

32. 45 C.F.R. § 46.103. The current version of the Federal Wide Assurance form is available at http://www.hhs.gov/ohrp/assurances/assurances/fwaformpdf.pdf (last visited May 7, 2013). As David Hyman points out, "This mismatch is non-trivial; . . . nearly 80 [percent] of all research projects reviewed by the University of Chicago's Social Science IRB are either personally funded, privately funded, or unfunded." Hyman, *supra* note 15, at 752.

33. Carol Weil et al., *OHRP Compliance Oversight Letters: An Update*, IRB: ETHICS & HUM. RES., Mar.–Apr. 2010, at 1, 5 (finding, based on "informal review of a sample of institutions," that in 2000, over 90% of domestic institutions had agreed to extend the regulations, compared to 74% in 2010); *see also* AM. ASS'N OF UNIV. PROFESSORS, INSTITUTIONAL REVIEW BOARDS AND SOCIAL SCIENCE RESEARCH 5 (2001), *available at* http://www.aaup.org/report/institutional-review-boards-and-social-science-research (about 75% "of the largest American research institutions" have voluntarily extended IRB review).

34. Human Subject Research Protections, 76 Fed. Reg. 44,512, 44,528 (July 26, 2011) [hereinafter ANPRM].

35. Judith Jarvis Thomson et al., *Research on Human Subjects: Academic Freedom and the Institutional Review Board*, ACADEME, Sept.–Oct. 2006, at 95, 99.

36. *See* INT'L COMM. OF MED. JOURNAL EDITORS, *Uniform Requirements for Manuscripts Submitted to Biomedical Journals: Writing and Editing for Biomedical Publication*, 6 (2010), *available at* http://www.icmje.org/urm_full.pdf. Of 103 major U.S., U.K., and Canadian biomedical journals, the proportion requiring IRB approval for publication increased from 42% in 1995 to 76% in 2005. A. Rowan-Legg et al., *A Comparison of Journal Instructions Regarding Institutional Review Board Approval and Conflict-of-Interest Disclosure Between 1995 and 2005*, 35 J. INST. MED. ETHICS 74, 75 (2009).

through her institution's contract with OHRP, her employment contract, or her publishing contract.

In addition, U.S. regulations have considerable global reach. They apply directly to research conducted or funded by a Common Rule agency that takes place outside the U.S. Foreign researchers who wish to market drugs, devices, or biologicals in the U.S., which is a leading consumer of these products, must comply with the FDA's essentially identical regulations, as do foreign researchers who wish to publish in many U.S. journals or who conduct HSR in one of the many countries that have modeled their own HSR protections on U.S. regulations.[37]

2. Covered Activities

The regulations thus cover a perhaps surprising number of actors. Due to a broad definition of "research," they apply to a similarly broad range of studies. For instance, IRBs review not only biomedicine and psychological research, but also research from virtually every social science, humanities, and professional discipline, including sociology, anthropology, history, economics, philosophy, memoir and biography, and classics. IRBs also review public policy "experiments" and research conducted in professional schools of law, business, education, and journalism. Additionally, they review research using virtually every methodology, from pharmacology and safety studies of investigational new drugs, to research on existing data and tissue, to surveys, interviews, and observation.

An activity is covered by the regulations if it (1) constitutes *research*: "a systematic investigation . . . designed to develop or contribute to generalizable knowledge;"[38] and (2) involves a *human subject*: "a living individual about whom an investigator (whether professional or student) . . . obtains [either] [d]ata through intervention or interaction with the individual, or . . . [i]dentifiable private information."[39] Formally, the Common Rule categorizes HSR into three levels of regulation: research that is subject to review by a fully convened IRB (the default), ten categories of research eligible for expedited IRB review, and six categories of HSR

37. *See* Maureen H. Fitzgerald et al., *The Research Ethics Review Process and Ethics Review Narratives*, 16 ETHICS & BEHAV. 377, 378–79 (2006) (observing "remarkable commonality" among Australian, Canadian, U.K., and U.S. review boards); Kevin D. Haggerty, *Ethics Creep: Governing Social Science Research in the Name of Ethics*, 27 QUALITATIVE SOC. 391, 393 (2004) (finding that U.S. and Canadian review boards are "comparable").

38. 45 C.F.R. § 46.102(d) (2012).

39. *Id.* § 46.102(f). *Intervention* "includes both physical procedures by which data are gathered . . . and manipulations of the subject or the subject's environment that are performed for research purposes," while *interaction* "includes communication or interpersonal contact." *Id.*

that are "exempt" from review.

In practice, however, the second and third of these levels tend to collapse into the first.[40] Although the regulations exempt six categories of research from IRB review,[41] they do not specify who determines whether a research proposal falls within one of these categories. In 1995, (then) OPRR issued guidance advising that "investigators should not have the authority" to make this decision and "should be cautioned to check with the IRB or other designated authorities."[42] By 1998, not surprisingly, nearly three-quarters of surveyed IRB administrators reported routine involvement in exemption determinations. And by 2003, most institutions had formally contracted with OHRP (via the FWA) to require researchers to submit protocols to the IRB to determine their exemption status.[43] Thus, most researchers must submit both exempt and non-exempt research to the IRB.

Moreover, because the regulations constitute a floor, not a ceiling,[44] even if an IRB determines that a protocol is exempt, it is not required to refrain from reviewing it. IRBs may—and regularly do—subject what are more accurately called exempt*ible* proposals to expedited or even full IRB review. The *Bell Report* found, for instance, that 15% of IRB-reviewed proposals were exemptible,[45] and that fewer than half of responding IRBs regularly exempted from review such exemptible research as analysis of existing data, interviews, and surveys.[46] Indeed, some IRBs, by policy, simply subject all protocols to full review.[47] As one commentator, himself an IRB member,

40. For a more detailed explanation of why IRBs tend to "define up" research into higher strata requiring more extensive review, thereby thwarting efforts at risk-based research regulation, see Michelle N. Meyer, *Three Challenges for Risk-Based (Research) Regulation: Heterogeneity Among Regulated Activities, Regulator Bias, and Stakeholder Heterogeneity*, *in* THE FUTURE OF HUMAN SUBJECTS RESEARCH REGULATION 8–10, 11 (I Glenn Cohen & Holly Fernandez Lynch eds., forthcoming 2014), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2218549.

41. *See* 45 C.F.R. § 46.101(b).

42. OFFICE FOR PROTECTION FROM RESEARCH RISKS (OPRR), OPRR REPS. 95-02: EXEMPT RESEARCH AND RESEARCH THAT MAY UNDERGO EXPEDITED REVIEW (May 5, 1995), *available at* http://www.hhs.gov/ohrp/policy/hsdc95-02.html.

43. NAT'L RESEARCH COUNCIL, PROTECTING PARTICIPANTS AND FACILITATING SOCIAL AND BEHAVIORAL SCIENCES RESEARCH, 71–78 (Constance F. Citro et al. eds., 2003).

44. 45 C.F.R. § 46.112 (providing that research "may be subject to further appropriate review and approval or disapproval" by institutional officials).

45. JAMES BELL ET AL., FINAL REPORT: EVALUATION OF NIH IMPLEMENTATION OF SECTION 491 OF THE PUBLIC HEALTH SERVICE ACT, MANDATING A PROGRAM OF PROTECTION FOR RESEARCH SUBJECTS 9 (1998).

46. *Id.* at 27–30.

47. *See* Michael J. Meehan & Marleina Thomas Davis, *Key Compliance Issues for*

put it: "There is no great gain in seeking [exempt] status . . . ."[48]

Those studies eligible for expedited review fare similarly. Proposed research that imposes "no more than minimal risk" on participants and also falls within one of ten categories specified by the Secretary of HHS is eligible for expedited review.[49] Under expedited review, the IRB chairperson or her designate can review the research proposal alone,[50] which is often, but not always, faster than full review. But, as with exemptible research, the IRB determines both whether proposed research falls within an expeditable category and whether it involves "no more than minimal risk." And, as with exemptible research, IRBs "may," but need not, expedite review of expeditable research.[51] As a result, much expedit*able* research, like much exempt*ible* research, receives full IRB review. The *Bell Report* found, for instance, that of those high-volume IRBs surveyed, only 52% regularly conducted expedited review of studies involving a simple blood draw, and only 60% did so for studies involving non-invasive data collection from adults.[52]

A final factor that contributes to the regulations' broad scope is the considerable vagueness of key regulatory language,[53] which, when

---

Institutional Review Boards, *in* CLINICAL RESEARCH LAW AND COMPLIANCE HANDBOOK 299, 309 (John E. Steiner, Jr. ed., 2006).

48. J. Michael Oakes, *Risks and Wrongs in Social Science Research: An Evaluator's Guide to the IRB*, 26 EVALUATION REV. 443, 457 (2002).

49. 45 C.F.R. § 46.110(a)–(b)(1) (2012).

50. *Id.* § 46.110(b). The reviewer may approve or require changes to a proposal, but must send the proposal to the full IRB for a determination that the proposal should be rejected. *Id.*

51. *Id.*

52. BELL ET AL., *supra* note 45, at 29–30, fig. 16.

53. For example, the Common Rule's definition of "research"—"a systematic investigation . . . designed to develop or contribute to generalizable knowledge," 45 C.F.R. § 46.102(d)—has caused considerable consternation among researchers, IRBs, and federal regulators about when investigations are sufficiently "systematic" and "generalizable." It is often unclear whether planning activities prefatory to a study, such as informal "piloting" of a survey instrument, might themselves constitute a "systematic investigation." NAT'L RESEARCH COUNCIL, *supra* note 43, at 147. Case studies also fall into a grey area with respect to whether they constitute a "systemic investigation." *See, e.g.*, UNIVERSITY OF MICHIGAN HUMAN RESEARCH PROTECTION PROGRAM, ACTIVITIES SUBJECT TO THE HRPP, OPERATIONS MANUAL—PART 4 (2012), *available at* http://www.hrpp.umich.edu/om/Part4.html (defining case studies as not exempt from IRB review, and, as such, not "systemic investigation").

As for generalizability, this criterion of "research" is both undefined in the *Belmont Report* and the Common Rule and yet is also the "cornerstone" of these moral and legal frameworks for regulating HSR. Tom L. Beauchamp & Yashar Saghai, *The Historical Foundations of the Research-Practice Distinction in Bioethics*, 33 THEORETICAL MED. & BIOETHICS 45, 52 (2012). When students conduct research primarily as a learning experience rather

combined with IRBs' risk aversion,[54] tends to lead IRBs to err on the side of more, rather than less, review.

II. THE HETEROGENEITY PROBLEM

> There is no reason that all human existences should be construed on some one, or some small number of patterns. . . . The same things which are helps to one person towards the cultivation of his higher nature, are hindrances to another. The same mode of life is a healthy excitement to one, . . . while to another it is a distracting burden . . . . Such are the differences among human beings in their sources of pleasure, their susceptibilities of pain, and the operation on them of different physical and moral agencies, that unless there is a corresponding diversity in their modes of life, they neither obtain their fair share of happiness, nor grow up to the mental, moral, and aesthetic stature of which their nature is capable.
>
> —John Stuart Mill, *On Liberty*[55]

In determining whether a research project's "[r]isks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result,"[56] an IRB must (1) determine the magnitude of research-related harms and benefits to participants as well as the value of the resulting knowledge; (2) discount these (dis)utilities by the probability that they will occur;[57] and (3) determine whether the resulting risks to participants are "reasonable in relation to" the project's aggregate expected benefits to participants and

---

than in an attempt to produce generalizable knowledge, is their work subject to IRB review? *See* NAT'L RESEARCH COUNCIL, *supra* note 43, at 147–48. Whether the lessons learned from quality improvement and quality assurance activities are sufficiently "generalizable" beyond the institutions in which they are conducted to bring these activities within the IRB system is also a perennial problem. *See* David Casarett, Jason H.T. Karlawish & Jeremy Sugarman, *Determining When Quality Improvement Initiatives Should Be Considered Research: Proposed Criteria and Potential Implications*, 283 J. AM. MED. ASS'N 2275 (2000).

Whether a study is "minimal risk" often plays a critical role in the kind of review it receives, and in some cases whether it is permissible at all, *see infra* notes 114–26 and accompanying text, yet the regulatory definition of "minimal risk" is notoriously ambiguous. NAT'L RESEARCH COUNCIL, *supra* note 43, at 32–34. Even knowing when a study involves "human subjects" is "not always straightforward." *Id.* at 149.

54. *See infra* Part III.

55. JOHN STUART MILL, *On Liberty, in* UTILITARIANISM 126, 197–98 (Mary Warnock ed., Fontana 1972) (1962).

56. 45 C.F.R. § 46.111(a)(2) (2012).

57. Technically, the regulatory language suggests that IRBs catalog *all* relevant harms and benefits to participants, without regard to their probability, and then discount their (dis)utility accordingly. In assessing the value of the knowledge to be produced, by contrast, IRBs are to employ a threshold of probability, counting 100% of the value only of that knowledge that is "reasonably expected" to result.

society. This Part argues that IRBs lack information about prospective participants' preferences necessary to determine these inputs.[58] Even if IRBs could overcome this *information problem*, prospective participant heterogeneity would present them with an *aggregation problem*; IRBs must make a single determination, applicable to all prospective participants, as to a study's risks, expected benefits, and the reasonableness of the ratio between these.

A. Heterogeneity in Research Risks

Research participants can and sometimes do suffer various psychosocial, economic, legal, and physical harms. But what amounts to a serious risk for one prospective participant will often pose a far more modest risk to a second prospective participant, and may even constitute an expected benefit for a third. Consider several common research-related risks.

1. Psychological Risks

a. Trauma Research and the Risk of Revictimization

Studies of sexual abuse and assault, grief, war, terrorism, natural disasters, and various other traumatic experiences are critical to gaining a better understanding of and addressing these phenomena. But exposure to trauma—whether as a survivor or as a first responder or other third party—often causes substantial psychological morbidity.[59] A meta-review of fifty-seven studies of natural disasters and their impacts on mental health, for instance, found that 74% of the victims sampled suffered post-traumatic stress and 39% were depressed.[60] Moreover, participants in trauma research may be struggling with medical, economic, or social difficulties secondary to the trauma.[61]

Given their potentially fragile state, IRBs understandably worry that "questioning [or otherwise studying] individuals who have experienced distressing events or who have been victimized in any number of ways . . . might rekindle disturbing memories, producing a form of re-

58. This Article considers only two of these inputs: risks and expected benefits to research participants. In future work, I plan to consider the third input—"the importance of the knowledge that may reasonably be expected to result" from research—which, I argue, involves a similar heterogeneity problem. 45 C.F.R. § 46.111(a)(2).

59. Sandro Galea et al., *Participant Reactions to Survey Research in the General Population After Terrorist Attacks*, 18 J. TRAUMATIC STRESS 461, 461 (2005).

60. FRAN H. NORRIS, RANGE, MAGNITUDE, AND DURATION OF THE EFFECTS OF DISASTERS ON MENTAL HEALTH: REVIEW UPDATE 2005 3 fig. 1 (2005).

61. Galea et al., *supra* note 59, at 461.

victimization."[62] In one proposed study, for instance, adults were to be asked to anonymously complete an online survey in which they would recall childhood memories of the death of a family member. Although this study was exemptible, the reviewing IRB member sent the application to full board review, finding that "subjects were at severe psychological risk of experiencing post-traumatic stress disorder."[63] The full IRB agreed, and required the researcher to provide participants with access to on-site psychological counseling. However, this meant that the online survey had to be administered locally, which severely limited the generalizability of the results.[64]

Revictimization and similar risks are sometimes dismissed by critics of IRBs as trivial, if not wholly imagined. IRB review, they say, should be reserved for biomedical research or studies that pose risks of physical harm.[65] Among the general population of trauma-exposed individuals, concerns about revictimization are not borne out; a majority of studies finds that trauma-exposed individuals do not experience severe or lasting distress associated with participation in trauma-focused research.

But there is little doubt that *some* individuals *will* fare worse for having recalled traumatic events. As such, these studies cannot be dismissed as per se "low-risk." Yet, while there is little doubt that participation will harm some significantly, there is equally little doubt that it will harm others only modestly, and benefit still others.

Consider one study involving three surveys of randomly selected residents of New York City conducted 1–2 months, 4–5 months, and 6–9 months after the September 11, 2001 terrorist attacks. Participants were asked about their exposure to the attacks and assessed for probable depression and post-traumatic stress disorder (PTSD). At the end of each interview, participants were asked whether they found any of the questions emotionally upsetting and, if so, whether they were still upset or were "okay now." Those who reported still being upset were asked if they would like a

62. Haggerty, *supra* note 37, at 400.

63. Nat'l Commc'n Ass'n, *Communication Scholars' Narratives of IRB Experiences*, 33 J. APPLIED COMM. RES. 204, 214 (2005) (anonymous narrative #18)*; see also* Wynn W. Gadar-Wilcox Comment to *IRBs and Clean Secrets*, A THAUMATURGICAL COMPENDIUM, (Dec. 9, 2011, 1:01 PM), http://alex.halavais.net/irbs-clean-secrets (IRB rejected study of war veterans' opinions of Bush Administration foreign policy "because of fears that the question might trigger [post-traumatic stress disorder (PTSD)]"); Oakes, *supra* note 48, at 446 (IRB questioned risk/benefit ratio of prisoner survey eliciting "memories of freedom").

64. In addition, requiring an on-site counselor and similar risk management techniques, as IRBs often do, entail nontrivial costs that effectively kill such research for researchers who lack outside funding, such as graduate students and many non-biomedical researchers.

65. These two categories are overlapping, but hardly coextensive.

counselor to call them. Of the 5,774 total participants surveyed in the three surveys, 13% said that the questions were upsetting, 1% were still upset at the end of the interview, and 0.3% were still upset and accepted the offer of counseling.[66] Those who were aged 45–64, female, single, lacked health insurance or a regular health care provider, were directly affected by the attacks, had current probable PTSD or depression or probable PTSD or depression since the attacks, or reported previous mental health problems in the year prior to the attacks were more likely to find the survey questions emotionally upsetting.[67] Significant participant heterogeneity remained even within these categories.[68]

Thus, although many scholars note that "[IRB] members differ on how they evaluate the seriousness of the harms associated with upsetting or traumatizing a research participant,"[69] as if this were evidence of errors in IRB risk–benefit analysis, heterogeneity in IRB risk assessment is more likely to reflect a significant degree of arbitrariness in IRB decisionmaking, given prospective participants themselves would likely differ on this question.

b. Sensitive Topics

Many studies ask participants to discuss socially disfavored behavior and other potentially "sensitive" topics such as drug use, gambling, risky or unconventional sexual behavior, HIV seropositivity, criminal behavior, and sexual assault and victimization. IRBs worry that such research will be emotionally arousing for participants, causing them embarrassment, fear, or general discomfort. As a result, a survey of the 450 members of the American Sociological Association Section on Sexualities found, "IRBs routinely block[] research on adult sexual minorities, particularly LGBTQ communities, because of their alleged vulnerability."[70] Of those who had submitted sexuality-related proposals to an IRB, 45% reported difficulty getting approval, and 41% reported that other sexuality researchers at their

66. Galea et al., *supra* note 59, at 461. Ninety-six percent of those who began the survey completed it. *Id.* at 462. In addition to the 0.3% of those who completed the survey and requested counseling (nineteen participants), ten participants "who were emotionally upset early in the interview" did not finish the survey, and received counseling. *Id.* at 463.

67. *Id.* at 463.

68. For instance, 45% and 27% of respondents with current probable PTSD or depression, respectively, reported that the survey questions were emotionally upsetting, compared to 12% and 11% of those deemed unlikely to currently suffer from these conditions. *Id.* at 464.

69. Haggerty, *supra* note 37, at 400.

70. Janice M. Irvine, *Can't Ask, Can't Tell: How Institutional Review Boards Keep Sex in the Closet*, CONTEXTS, Spring 2012, at 28, 30, 32 (response rate: around 40%).

university had also had IRB difficulties. Some were merely slowed down in their research, while others yielded to conditions that reduced the value of that research, such as IRB demands that interview tapes be destroyed, which precludes longitudinal follow-ups and use by future historians. Still others reported abandoning research on these topics and counseling students to do likewise.[71] IRBs have responded similarly to other studies involving sensitive topics.[72]

Again, however, these IRB concerns are not unfounded. One survey asked two groups of men—those in the general population and men who have sex with men (MSM)—to report their level of discomfort after being asked questions about illicit drug use. Although the mean level of discomfort reported by both groups was relatively low (1.78 and 1.66 out of 7 for the general population and MSM samples, respectively), some respondents reported greater discomfort than others.[73] Non-white respondents and those who reported having used illicit drugs within the past year, for instance, reported more discomfort than did white respondents and those who did not report having used illicit drugs within the prior year.[74] Similarly, when researchers asked female undergraduates about various sensitive topics, those who had experienced child abuse were more likely to report distress due to remembering the past than were other respondents.[75]

Again, such participant heterogeneity may be reflected in the lack of consensus among IRBs regarding the level of risk posed by sensitive topics.

---

71. *Id.* at 30.

72. In one study, an IRB effectively forced an undergraduate under a graduation deadline conducting survey research for her thesis to abandon a question on undergraduates' views of reparations after the IRB decided the study required full review due to the "sensitive" nature of that question. Ross Cheit, Comment to *Outside of Biomed Research, IRBs are Essentially Censorship Agencies*, EMPIRICAL LEGAL STUD. BLOG (July 9, 2006, 9:25 PM), http://www.elsblog.org/the_empirical_legal_studi/2006/03/outside_of_biom.html. Another IRB prohibited a Caucasian graduate student from asking African-American graduate students about their career expectations for fear that the experience might "be traumatic" for them. Thomson et al., *supra* note 35, at 96. Indeed, researchers who submitted to different IRBs proposals that were identical except for the political significance of the propositions they proposed to test found that IRB decisions varied considerably depending on the presence or absence of political controversy. Stephen J. Ceci et al., *Human Subjects Review, Personal Values, and the Regulation of Social Science Research*, 40 AM. PSYCHOLOGIST 994, 994–95 (1985).

73. *See* Michael Fendrich et al., *Respondent Reactions to Sensitive Questions*, J. EMPIRICAL RES. ON HUM. RES. ETHICS, Sept. 2007, at 31, 32 & tbl.1.

74. *Id.* at 32.

75. Suzanne E. Decker et al., *Ethical Issues in Research on Sensitive Topics: Participants' Experiences of Distress and Benefit*, J. EMPIRICAL RES. ON HUM. RES. ETHICS, Sept. 2011, at 55, 55.

A survey of 188 randomly selected IRB chairpersons found that while 44% considered a confidential survey of healthy eleven-year-olds about sexual behavior to pose minimal risk, 29% considered it a minor increase over minimal risk, and 19% considered it more than a minor increase over minimal risk.[76] Under federal regulations governing research with minors,[77] research that does not "hold out the prospect of direct benefit for the individual" participant[78] and is deemed to pose more than minimal risk is usually unapprovable under HHS regulations.

c. Unpleasant Self-Knowledge

Or consider research in which participants may learn something unpleasant about themselves. In order to test the hypothesis that children of alcoholics who are resilient are less likely to become alcoholics than are those who are less resilient, a researcher proposed to survey college students about their alcohol use and measure their resiliency. Although the consent form identified the risk that participants might learn that they may have a drinking problem and provided participants with referral information, the IRB rejected the study because of this risk.[79]

A college student may well respond negatively to this information, or may reap a net benefit. The information might lead him to pursue formal treatment, obtain a second opinion, increase his self-monitoring, or limit his drinking. Some research even suggests that those with relatively little resilience can increase it through deliberate effort.[80] Conversely, of those participants who score "normally" on the alcoholism screening test, some may benefit from being relieved of a fear that their family history destined them to alcoholism, while others may gain a false sense of security that

76. Seema Shah et al., *How Do Institutional Review Boards Apply the Federal Risk and Benefit Standards for Pediatric Research?*, 291 J. AM. MED. ASS'N 476, 479 & tbl.2 (2004).

77. *See infra* note 114.

78. Such research is approvable under HHS regulations only if it, *inter alia*, poses "a minor increase over minimal risk" and is "likely to yield generalizable knowledge about the subject's disorder or condition," 45 C.F.R. § 46.406 (2012), or, *inter alia*, the IRB finds that it "presents a reasonable opportunity to further the understanding, prevention, or alleviation of a serious problem affecting the health or welfare of children" and a panel specially convened by the Secretary approves it after opportunity for public review and comment. *Id.* § 46.407. Federal regulations governing research with prisoners, see 45 C.F.R. 46 subpart C, are even more stringent, and typically permit such research only if it poses "no more than minimal risk and no more than inconvenience to the subjects." *Id.* § 46.306.

79. *Telling People What They Already Know Can Hurt Them*, IRBWATCH (on file with author).

80. *See* DENNIS S. CHARNEY & CHARLES B. NEMEROFF, THE PEACE OF MIND PRESCRIPTION 18 (2004) (citing research from multiple disciplines showing that resilience can be improved with intentional effort).

emboldens them to abuse alcohol.

In another study, a researcher proposed anonymously surveying college students about their experiences of what the consent materials called "sexual aggression victimization." She deliberately avoided the term "rape" because "*inform*[ing] respondents currently unaware of their rape victim status that indeed they are rape victims may actually instigate trauma rather than prevent it." She also argued that the alteration "could significantly diminish the validity of the results" because "[m]any rape victims may refuse to admit to their victim status, thereby excluding themselves from the study altogether." As a result, the study would "fail to produce findings of potential assistance to the very victims the IRB apparently wants to assist." But the IRB disagreed and required her to refer specifically to "rape" because (in the researcher's words) "victims of rape 'need to know' they are victims of rape."[81] As with the alcoholism study, however, different participants will likely have a wide range of reactions to learning that they are rape victims.

The risk of potentially unpleasant self-knowledge is also posed by a growing body of research that many legal academics deem critical to issues as varied as employment discrimination and affirmative action,[82] legal decisionmaking,[83] and health disparities.[84] Implicit bias research uses an interactive, computer-based test in which participants quickly categorize words or images that appear on the screen by pressing a key corresponding to a given category.[85] Millions of people have participated in this research online,[86] and most were told that, whatever they might previously have thought, their "data suggest" that they harbor biased associations about people on the basis of sex, race, or other categories. Like other forms of

81. Nat'l Commc'n Ass'n, *supra* note 63, at 208–09 (anonymous narrative #7).

82. *See* Samuel R. Bagenstos, *Implicit Bias, "Science," and Antidiscrimination Law*, 1 HARV. L. & POL'Y REV. 477, 477–78 (2007) (citing empirical research on implicit bias).

83. *See* Jeffrey J. Rachlinski et al., *Does Unconscious Racial Bias Affect Trial Judges?*, 84 NOTRE DAME L. REV. 1195, 1197 (2009) (finding that trial judges hold implicit racial biases that may influence their judgment); Theodore Eisenberg & Sheri Lynn Johnson, *Implicit Racial Attitudes of Death Penalty Lawyers*, 53 DEPAUL L. REV. 1539, 1542 (2004) (suggesting that defense counsel hold implicit racial biases); Reshma M. Saujani, *"The Implicit Association Test": A Measure of Unconscious Racism in Legislative Decision-Making*, 8 MICH. J. RACE & L. 395, 396 (2003) (arguing that the current legal discrimination framework does not reach unconscious discrimination).

84. *See* Michael S. Shin, *Redressing Wounds: Finding a Legal Framework to Remedy Racial Disparities in Medical Care*, 90 CALIF. L. REV. 2047, 2049 (2002) (asserting that implicit racial bias is a leading cause of medical treatment disparities).

85. This description is based on the Implicit Association Test (IAT), which is the basis for most implicit bias research cited in the legal academic literature. *Id.* at 2066–67.

86. *See, e.g.*, Rachlinski et. al., *supra* note 83, at 1198 ("More than four and a half million people have taken the IAT.").

risky self-knowledge, learning that one is implicitly biased is likely to produce negative emotions in many participants. They may feel significant shame and powerlessness to change or otherwise respond constructively. Others, however, may dismiss the results as pseudo-science or the result of a bad test day. Still others may have a net positive reaction: though unsettling, the results may prompt them to learn more about implicit bias, to try to debias themselves through increased contact with "the other,"[87] or to rethink their positions on issues like affirmative action.[88]

2. Informational Privacy Risks

Research often poses risks of personally identifiable information being disclosed. Although estimating the likelihood of inadvertent disclosure will typically be a matter of technical expertise—albeit not the variety that IRBs usually possess[89]—the likelihood that disclosure will harm the participant and the magnitude of that harm depend on individual preferences and circumstances.

Individuals differ widely in their attitudes toward informational privacy. Researchers have found that these differences correlate with gender,[90] age, and extensiveness of social media use.[91] Privacy preferences are even heterogeneous *within* individuals; like many other preferences, they tend to be unstable or context-dependent. For instance, individuals' privacy concerns tend to decrease with more experience.[92] They may also vary depending on the kind of personal information at issue,[93] the perceived

87. *See, e.g.*, Christine Jolls & Cass R. Sunstein, *Debiasing Through Law*, 35 J. LEGAL STUD. 199, 201–04 (2006) (asserting that debiasing through law could allow people to change their own behavior).

88. Another example of self-knowledge that gives regulators and commentators pause is the individual results of research, especially genetic research. *See generally* NAT'L INSTS. OF HEALTH, GENOME-WIDE ASSOCIATION STUDIES (GWAS): NIH POINTS TO CONSIDER (2011), gwas.nih.gov/pdf/PTC_for_IRBs_and_Institutions_revised5-31-11.pdf.

89. ANPRM, *supra* note 34, at 44,516. For this reason, one proposed amendment to the regulations would take assessment of informational risks away from IRBs and require all researchers to comply with Health Insurance Portability and Accountability Act data-security standards. *Id.* at 44,515.

90. For instance, not surprisingly, girls are more concerned than boys about disclosing information pertaining to their physical location. *See* Ian Brown, Privacy Attitudes, Incentives and Behaviours 2–3 (June 17, 2011) (unpublished manuscript), *available at* http://ssrn.com/abstract=1866299.

91. *See* Alice E. Marwick et al., *Youth, Privacy and Reputation* 4, 7, 65 (Harvard Law Sch. Pub. Law & Legal Theory Working Paper Series, Paper No. 10-29, 2010), *available at* http://ssrn.com/abstract=1588163.

92. Brown, *supra* note 90, at 3–4.

93. *See* Steve Jones et al., *Everyday Life, Online: U.S. College Students' Use of the Internet*, FIRST MONDAY (Oct. 5, 2009), http://firstmonday.org/ojs/index.php/fm/article/view/2649/

tradeoffs involved,[94] the intended use of the information,[95] the voluntariness of the disclosure, and the perceived trustworthiness of the recipient.

For one dramatic example of individual differences in privacy preferences, consider the Personal Genome Project (PGP), run by Harvard geneticist George Church. Many view genetic and medical information as among the most private kinds of information that exist. Yet, participants in the PGP agree to have their entire genome sequenced and, along with detailed medical and other personal information, *posted on the Internet* for anyone to see, download, and analyze. The first ten participants are identified by name and photograph, and the profiles of most other participants are so rich that they can be easily identified through data mining.[96]

3. Physical Risks

To date, there has been only one serious attempt to develop a disciplined method of assessing research risks that improves upon IRB (and regulator) intuition. The Systematic Evaluation of Research Risks (SERR) limits itself to physical risks "due to the[] strong context dependence" of economic and social harms.[97] Indeed, those who would concede participant heterogeneity in psychological and social risks may be more skeptical of the existence of significant heterogeneity in seemingly more objective physical risks.[98] But,

---

2301 (finding few students at forty U.S. colleges (n=7,421) were concerned about disclosure of personal information on social networking sites but nearly 75% concerned about security of passwords and social security and credit card numbers).

94. *See, e.g.*, Seounmi Youn, *Teenagers' Perceptions of Online Privacy and Coping Behaviors: A Risk–Benefit Appraisal Approach*, 49 J. BROADCASTING & ELECTRONIC MEDIA 86, 100–01 (2005).

95. *See, e.g.*, Havasupai Tribe v. Ariz. Bd. of Regents, 204 P.3d 1063, 1066–67 (Ariz. Ct. App. 2008) (members of a Native American tribe who provided tissue samples to researchers to study the tribe's diabetes epidemic objected when researchers investigated tribal ancestry and prevalence of schizophrenia and inbreeding); Anne Adams & Martina Angela Sasse, *Privacy in Multimedia Communications: Protecting Users, Not Just Data*, *in* 49 PEOPLE AND COMPUTER XV—INTERACTION WITHOUT FRONTIERS: JOINT PROCEEDINGS OF HCI2001 AND ICM2001 49, 57 (A. Blandford et al. eds., 2001) (describing how some individuals were willing to permit videoconferencing recordings to be shared for purposes of evaluating the technology, but not in order to evaluate the technology's effects on different ethnic groups).

96. Participants' profiles are available at http://www.personalgenomes.org/consent/whitepaper_consent_04302007.pdf. The first ten participants were required to hold an M.A. in genetics or its equivalent. The detailed informed consent process for later participants includes a requirement that enrollees receive a perfect score on a genetics exam.

97. Annette Rid et al., *Evaluating the Risks of Clinical Research*, 304 J. AM. MED. ASS'N 1472, 1473–74 (2010).

98. Legal and economic risks, too, might seem to be the province of technocrats. For

in fact, physical risks—no less than psychological, social, legal, and economic risks—depend on preferences and other individual circumstances. Despite a clear role for expertise in assessing physical risks, then, much critical information regarding these risks, too, remains privately held by prospective participants.[99]

a. Pain Heterogeneity

Of the seven dimensions that SERR uses to assess the riskiness of research, the first—"experience, such as pain, associated with the harm"[100]—would seem to vary the least among individuals. Yet "[o]ne of the most striking features of pain is the large range of variation in response to identical stimuli."[101] The magnitude of pain caused by an identical stimulus can vary within the same individual over time.[102] It can also vary among individuals. The same injury or disease process, for instance, can result in chronic pain for one individual but minimal deficits for another

---

an argument that legal decisions mirror medical treatment (and research participation) decisions in combining expert and individual lay knowledge, see Note, *The Plaintiff as Person: Cause Lawyering, Human Subject Research, and the Secret Agent Problem*, 119 HARV. L. REV. 1510, 1511–12 (2006).

99. Bioethics emerged as a field during the anti-authoritarian 1960s, largely in opposition to medicine's paternalistic Hippocratic tradition. Bioethicists rightly pointed out that although patients rely on experts to tell them the relative "success" of, say, mastectomy versus lumpectomy with radiation in shrinking tumors, they rely on their own knowledge of their values, preferences, and circumstances to decide which of these (or neither) is most likely to be "successful" for them within the broader context of their lives. Given a choice between mastectomy or lumpectomy with radiation, some women may choose the latter because they feel that their breasts are integral to their identity or because they value the experience or option of breastfeeding children, while others may choose the former if it carries even a small relative increase in life expectancy. Bioethicists should be the last to express surprise, then, that IRBs vary markedly in their assessments not only of psychosocial risks but also of physical risks, and it is ironic that this lesson seems to have been largely forgotten in the domain of research. For a history of bioethics, see generally Daniel Callahan, *Bioethics and Policy—A History*, *in* FROM BIRTH TO DEATH AND BENCH TO CLINIC: THE HASTINGS CENTER BIOETHICS BRIEFING BOOK FOR JOURNALISTS, POLICYMAKERS, AND CAMPAIGNS ix (Mary Crowley ed., 2008), *available at* http://www.thehastingscenter.org/Publications/BriefingBook/Detail.aspx?id=2412.

100. Rid et al., *supra* note 97, at 1473.

101. Christopher S. Nielsen et al., *Characterizing Individual Differences in Heat-Pain Sensitivity*, 119 PAIN 65, 65 (2005).

102. An individual's subjective experience of pain can vary substantially from day to day, despite being evoked by an identical stimulus. *See* Robert C. Coghill et al., *Neural Correlates of Interindividual Differences in the Subjective Experience of Pain*, 100 PNAS 8538, 8538 (2003). This is likely to due to modulating factors such as anxiety. *See* Allan Jones et al., *Dispositional Anxiety and the Experience of Pain: Gender-Specific Effects*, 7 EURO. J. PAIN 387, 388, 393 (2003).

(interindividual differences), likely due to a combination of genetic and environmental factors.[103]

It is not uncommon to hear individuals describe themselves as particularly sensitive to, or tolerant of, pain. Studies have lent credence to such statements, finding that individuals' subjective pain ratings correlate with activity levels in the relevant areas of the brain,[104] thereby providing "crucial evidence that individual differences in *reported* pain reflect actual differences in *experienced* pain."[105]

Researchers have found individual differences in how individuals perceive both the intensity and the unpleasantness of pain to be "remarkably large."[106] Indeed, pain ratings of identical noxious stimuli can cover the entire scale from "no pain" to "the most intense pain [imaginable]."[107] And in heat pain and cold-pressure pain studies, researchers have found "no temperature that is painful to all subjects and at the same time tolerable to all subjects."[108]

In one small study that used brain imaging to try to identify objective neural correlates of subjective experiences of pain, individuals' reports of pain intensity evoked by the same 49°C noxious stimulus delivered to each participant's lower right leg ranged, on a ten-point scale, from 1.05 to 8.9. Moreover, the distribution of these results was remarkably uniform: rather than forming a bell curve of typical individuals, with relatively pain-sensitive and pain-insensitive outliers on each end, the scatter plot "curve" was actually a straight, diagonal line. In other words, the odds of randomly selecting a pain-sensitive and pain-insensitive participant are about the same, according to the study.[109]

In another study, investigators subjected 175 healthy participants to tests of heat-induced pain in both ascending and random series over three and one-half hours. Immediately following each stimulus, participants were asked to rate both its pain intensity and its discomfort on a scale that ranged from none to the worst intensity or discomfort they could imagine. Their ratings were then converted to a scale of 1 to 100, as is conventional in such studies. Investigators found that individual differences in pain

---

103. Nielsen et al., *supra* note 101, at 66; *see also* Amanda C. Pustilnik, *Pain as Fact and Heuristic: How Pain Neuroimaging Illuminates Moral Dimensions of Law*, 97 CORNELL L. REV. 801, 810–11 (2012) (pain is "inherently variable, subjective, and individual").

104. Coghill et al., *supra* note 102, at 8538, 8541.

105. Nielsen et al., *supra* note 101, at 66.

106. *Id.* at 68; *see also* Christopher S. Nielsen et al., *Individual Differences in Pain Sensitivity: Genetic and Environmental Contributions*, 136 PAIN 21, 27 (2008).

107. Nielsen et al., *supra* note 101, at 66.

108. *Id.* at 73.

109. Coghill et al., *supra* note 102, at 8539 & fig. 1.

sensitivity accounted for more of the total variance in the study (60%) than did the different temperatures of various stimuli themselves (40%).[110] Moreover, these individual differences in pain sensitivity and intensity are not stable or, put another way, "pain" is not a monolithic phenomenon. Investigators estimated that genetics accounts for 26% of the individual differences in sensitivity to heat pain, and researchers found no correlation with gender. When the same researchers investigated cold-pressor pain, however, although they found a similarly large amount of individual variation in pain sensitivity, they estimated that genetics accounted for 60% of this variation, and reported "significant" gender differences, with women reporting more pain than men.[111]

b. Heterogeneity in Other Aspects of Physical Risk

It is even easier to see that the remaining aspects of harm SERR uses to classify degrees of risk—the "burden of efforts, including treatment, to mitigate the harm," the "effects on an individual's ability to perform the activities of daily life" and to "pursue life goals," the "extent to which an individual can adapt to the new circumstances," and the "burden imposed by the process of adaptation"—will also vary considerably among prospective participants.[112] This is because even when individuals experience the same degree of pain or disability, identical hedonic experiences often have different *meanings* and *consequences* for different individuals (or for the same individual at different times). Some kinds of physical harms will be more significant for some than for others.[113] For athletes and musicians, for instance, physical disability may be significantly more harmful (where harm entails a setting back of one's interests) than pain, whereas for lawyers, scholars, business executives, and others who must be able to think clearly, pain or cognitive or psychological disabilities might be far worse than many physical disabilities. Even the same injury may have markedly different effects on the ability of different people to pursue *their* life goals and to perform *their* daily life activities. The efforts involved in treating that injury may be more or less burdensome, depending on each individual's access to treatment. And should the injury result in permanent disability, individuals may differ in the extent to which they can adapt to this new circumstance and, if so, in the burdens they would bear in so doing.

110. Nielsen et al., *supra* note 101, at 74.

111. Nielsen et al., *supra* note 106, at 25–26.

112. The seventh aspect of Systematic Evaluation of Research Risks (SERR) is duration of harm. Rid et al., *supra* note 97, at 1473.

113. *See supra* note 99.

It therefore should not be surprising that IRBs vary widely in their assessments of the riskiness of studies and common research procedures. IRBs vary in applying the regulatory distinctions, critical in the U.S.[114] and other jurisdictions,[115] among research that entails "minimal risk," research that entails a "minor increase over minimal risk," and research that entails more than a minor increase over minimal risk.[116] In one study, 188 IRB chairs were asked to determine into which of these three regulatory categories several procedures routinely used in biomedical research fall, assuming that participants were healthy eleven-year-olds. The results revealed "substantial"—and, according to the study authors, "unjustified"—variability in risk assessment. Twenty-three percent of chairs categorized allergy skin testing as minimal risk, while 43% categorized it as a minor increase over minimal risk and 27% categorized it as more than a minor increase over minimal risk.[117] On the other hand, most (81%) thought a one-time blood draw constituted minimal risk,[118] but one can easily imagine that for many, genetic privacy concerns would render such a procedure "risky."

Another study of IRB chairs in Germany similarly found a "disturbingly high degree" of variation in the assessment of physical risks. Twenty-nine

---

114. In the U.S., whether research poses "no more than minimal risk" determines whether it is eligible for expedited review, whether minors may participate in nontherapeutic research, and whether informed consent requirements may be altered or waived. *See* 45 C.F.R. § 46.110 (2012) (expedited review); *id.* §§ 46.116(a)(6), (d)(1), 46.117(c)(2) (informed consent); *id.* § 46.204(b), (d) (pregnant women and fetuses); *id.* § 46.306(a)(2)(i), (ii) (prisoners); *id.* § 46.404 (children). "Minimal risk" research is defined for most of these purposes as research whose anticipated "probability and magnitude of harm or discomfort . . . are not greater in and of themselves than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests." *Id.* § 46.102(i). The regulations define minimal risk slightly differently—expressly adopting an absolute rather than relative standard—in the context of research with prisoners. *See id.* § 46.303(d) ("Minimal risk is the probability and magnitude of physical or psychological harm that is normally encountered in the daily lives, or in the routine medical, dental, or psychological examination of healthy persons.").

115. *See* Loretta M. Kopelman, *Minimal Risk as an International Ethical Standard in Research*, 29 J. MED. & PHIL. 351 (2004) (explaining the role of "minimal risk" in other countries and in international codes and guidelines).

116. HHS regulations governing research with minors adds to "minimal risk" research the categories of research that involves a "minor increase over minimal risk" and, by implication, research that involves more than a minor increase over minimal risk, neither of which the regulations or HHS further defines. 45 C.F.R. § 46.406.

117. *See* Shah et al., *supra* note 76, at 476.

118. *Id.* at 476, 478. The percentage of IRB chairs who classified the following procedures as minimal risk were as follows: sex surveys (44%), MRI (48%), one blood draw per week for 24 weeks (15%), electromyogram (8.5%), pharmacokinetic testing (7.5%), lumbar punctures (2%). *Id.* at 479 tbl. 2.

chairs were given five hypothetical, but realistic, biomedical studies involving children and asked to state for each whether they would approve it with or without "restrictions," would approve it "under no circumstances," or were unsure what they would do.[119] The authors intentionally chose the five protocols to reflect a range of risk from least risky (Study 1) to most risky (Study 5).

The chairs' decisions regarding what the authors describe as the three "highest-risk" protocols varied markedly. Study 4, for instance, called for an additional six bone marrow biopsies in leukemia patients already receiving four biopsies for therapeutic purposes. About 58% of chairs said they would not permit the study under any circumstances, with one characterizing the additional biopsies as "a kind of child abuse."[120] But 41% of chairs would have approved the study—half without restrictions.[121] Similarly, 48% of chairs would have approved Study 5, supposedly the riskiest study—again, half of them without restrictions—while 41% would have rejected it, and the remaining 10% or so were uncertain.[122]

This variation is not likely explained by disparities in the chairs' experience or expertise, however measured.[123] Just under 90% were physicians. Although only about 20% specialized in pediatrics, 93% had children of their own. About three-quarters had participated in more than fifty committee meetings, while the remainder had participated in ten to fifty meetings.[124]

Moreover, although there was broad consensus among the chairs about the acceptability of the two "least risky" protocols (all chairs would have approved both studies, although some would have required "restrictions" in one or both cases), this consensus contrasts sharply with the attitudes of the German public. As the authors note, Study 1, the "least risky" of the five, "strongly resembled" a 1997 case that sparked considerable outrage among Germans. In that case, a doctoral candidate had taken blood samples from residents of a home for the "mentally handicapped," for the purpose of doing genetic research. Although he informed neither the residents' guardians nor the residents themselves about his research, public criticism

119. C. Lenk et al., *Non-Therapeutic Research with Minors: How Do Chairpersons of German Research Ethics Committees Decide?*, 30 J. INST. MED. ETHICS 85, 86 (2004).

120. *Id.*

121. *Id.*

122. *Id.* at 86–87. Study 3 yielded similarly divergent results, with a full 20% of chairs unable to make any decision. *Id.* at 86 fig. 1.

123. Except, of course, as measured by the chairs' knowledge, or lack thereof, regarding the preferences of hypothetical individual research participants and, in this case, their proxy decisionmakers.

124. Lenk et al., *supra* note 119, at 86.

focused "predominantly[] on the alleged immorality of research without potential direct benefit."[125] Largely on the basis of this case, Germans successfully objected to their government ratifying the European Convention on Human Rights and Biomedicine, which permits minimal risk but nontherapeutic research on individuals unable to give consent.[126] Study 1 similarly proposed that researchers draw a small additional amount of blood from children who were already undergoing blood draws for therapeutic purposes. All twenty-nine chairs said they would approve the study, and only one would have required restrictions.[127]

B. Heterogeneity in Research Benefits

Participants choose to enroll in research for a range of reasons as broad as the range of risks they thereby assume.

1. Altruism and Pro-Sociality

Although neoclassical economics models individual behavior as motivated by the self-interested pursuit of extrinsic, material benefits,[128] an ample literature from multiple disciplines,[129] as well as data tracking donations of time and money,[130] suggests that human beings are motivated

125. *Id.* at 85–86.

126. *Id.* at 85.

127. *Id.* at 86.

128. *See, e.g.*, Oliver E. Williamson, *Transaction Cost Economics: How It Works; Where It is Headed*, 146 DE ECONOMIST 23, 31 (1998); William T. Allen, *Contracts and Communities in Corporation Law*, 50 WASH. & LEE L. REV. 1395, 1401 (1993).

129. *See, e.g.*, Yochai Benkler, *The Unselfish Gene*, HARV. BUS. REV., July–Aug. 2011, at 77, 77 (discussing how fields as diverse as evolutionary biology, psychology, sociology, political science, and experimental economics "are tracing a new intellectual arc in the disciplines concerned with human action and motivation" that undercuts the "deep-rooted belief about human selfishness"); *see also* YOCHAI BENKLER, THE PENGUIN AND THE LEVIATHAN: HOW COOPERATION TRIUMPHS OVER SELF-INTEREST (2011); Colin Camerer & Richard H. Thaler, *Anomalies: Ultimatums, Dictators and Manners*, J. ECON. PERSP., Spring 1995, at 209 (reporting evidence of other-regarding behavior from experiments with ultimatum and dictator games).

130. In the U.S., 93 million volunteers donate more than 20.3 billion hours every year to nonprofit organizations. Mary Bridgman, *Volunteers Answer Call Without Calling Attention to Themselves*, COLUMBUS DISPATCH, Nov. 29, 1998, at 1G, 8G. Theories of the determinants of volunteering for nonprofit organizations "are so varied and contradictory that no single conceptual model has received general support." Janet C. Winniford et al., *Motivations of College Student Volunteers: A Review*, 34 J. STUDENT AFF. RES. & PRAC. 135 (1997). In 2009, charitable giving in the U.S. totaled $303.75 billion, 75% of which came from individuals (or 88%, counting charitable bequests and estimated family foundation grants). Total giving has increased in current dollars every year but in 1987 and 2009. Between 1969 and 2009, annual total giving ranged from 1.7% to 2.2% of GDP, with 90% of people giving money to

by many factors, including those that are intrinsic, intangible, other-regarding and duty-driven. We are, in other words, only "boundedly self-interested."[131]

Much of the utility of altruistic participation in research may be already accounted for when IRBs weigh "the importance of the knowledge that may reasonably be expected to result"[132] from a study. But evidence suggests that those who engage in prosocial activities, including serving as research participants, often themselves thereby receive "warm glow utility."[133] Therapists successfully advise those who are grieving or depressed to take their minds off their own problems by focusing on the problems of others. Elderly individuals who volunteer report greater quality of life than those who do not volunteer.[134] And a large and growing body of empirical research finds strong associations between prosocial behavior and mental and physical health and well-being.[135] In one study, for example, about half of participants who helped others in modest ways—for example, by working at a soup kitchen for a few hours—reported a feeling of elation, or a sense of significance or meaning in life, while 13% reported a reduction in their chronic aches and pains.[136] Another study suggests a possible physiological basis for warm glow utility. Subjects positioned in a functional magnetic resonance imaging (fMRI) machine were asked to contemplate a menu of possible charities to which they might like to contribute. Researchers found that when subjects checked a box next to their preferred charity, the mesolimbic pathway, which is associated with dopamine, was activated.[137]

---

at least one charity. CTR. ON PHILANTHROPY AT IND. UNIV., GIVING USA 2010: THE ANNUAL REPORT ON PHILANTHROPY FOR THE YEAR 2009 5, 11 (2009).

131. *See* Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1545 (1998) (people's utility function includes concern "about the well-being of others, even strangers in some circumstances").

132. 45 CFR 46.111(a)(2) (2012).

133. *See* Peter A. Diamond & Jerry A. Hausman, *On Contingent Valuation Measurement of Nonuse Values*, *in* CONTINGENT VALUATION: A CRITICAL ASSESSMENT 3, 27 (Jerry A. Hausman ed., 1993) (describing warm glow valuation as the moral satisfaction individuals obtain when supporting good causes).

134. Morris A. Okun & Josef Michel, *Sense of Community and Being a Volunteer Among the Young-Old*, 25 J. APPLIED GERONTOLOGY 173, 186 (2006).

135. *See* Stephen G. Post, *Altruism, Happiness, and Health: It's Good to Be Good*, 12 INT'L J. BEHAV. MED. 66 (2005).

136. Allan Luks, *Helper's High*, PSYCHOL. TODAY, Oct. 1988, at 39.

137. Jorge Moll et al., *Human Fronto-Mesolimbic Networks Guide Decisions About Charitable Donation*, 103 PNAS 15,623 (2006). The researchers speculated that since merely contemplating money donation triggered this effect, actual donation of money (or other resources) would trigger an effect at least as great. However, it is also possible that individuals donating actual money (or other resources) might experience disutility associated

But warm glow utility is difficult to measure—especially in a metric that renders it commensurable with other utilities and disutilities and allows meaningful trade-offs to be made among these.[138] More to the present point, even if participants' warm glow utility could be measured and monetized, not all participants can be expected to receive equal, or any, warm glow utility. Individuals are heterogeneous in their preferences for prosociality.[139] Experiments testing cooperative behavior have found that while about 30% of individuals behave more or less like *homo economicus*, more of them—some 50%—behave cooperatively, either contingent on another's cooperative behavior or unconditionally. The remaining 20% of individuals behave unpredictably, sometimes cooperating and sometimes not.[140] IRBs, of course, have no method of discerning which prospective subjects would derive warm glow utility from their participation and which would not.

2. Compensation

Research participants are sometimes motivated, in full or in part, by money they receive. Some participate only once or occasionally to secure pocket change. For others, participating in research constitutes part- or even full-time employment.[141] Needless to say, the same amount of compensation will be more or less attractive to different prospective participants according to such individualized factors as their socioeconomic status, whether they currently have any pressing need for money, their taste for expensive goods, and so on.

In one study of participant perceptions of research compensation, sixty individuals who had previously received payment for participating as a

with that sacrifice that is equal to or greater than this warm glow utility.

138. *Cf.* Matthew D. Adler & Eric A. Posner, *Rethinking Cost-Benefit Analysis*, 109 YALE L.J. 165, 175 (1999) ("[Cost–benefit analysis (CBA)] is frequently hampered by a lack of data or by the difficulty of estimating valuations. A striking example is a CBA that attempted to monetize the aesthetic value that people attach to clear air over the Grand Canyon.").

139. Similarly, there is evidence of individual differences in risk perception along dimensions of socio-demographics, religion, general trust level, cultural factors, personal facets, and experience or information learning process. *See* Jan Urban & Milan Ščasný, Determinants of Risk Perception Bias: An Empirical Study of Economically Active Population of the CR, Paper presented at the "World of Labour and Quality of Life in Globalized Economy" Conference at the University of Economics at Prague 8 (Sept. 2007).

140. Benkler, *supra* note 129, at 79.

141. *See* ROBERTO ABADIE, THE PROFESSIONAL GUINEA PIG: BIG PHARMA AND THE RISKY WORLD OF HUMAN SUBJECTS 5 (2010) (ethnography of self-defined "professional guinea pigs" who earn around $15,000 to $20,000 (in 2002 dollars) by participating as healthy controls in five to eight drug trials per year); James A. Anderson & Charles Weijer, *The Research Subject as Wage Earner*, 23 THEORETICAL MED. 359, 361 (2002).

healthy volunteer in at least one clinical trial were asked to state the appropriate amount that participants should be offered in each of four hypothetical clinical trials, and whether they would be willing to participate. The monetary amounts given varied more from participant to participant than it did from hypothetical to hypothetical. That is, respondents had individualized, but consistent methods of arriving at estimates of payments for participating in clinical studies based on each individual's perception of study burden and associated risk.[142]

3. Beyond "For Love or Money": Other Benefits

Research can benefit participants in myriad ways beyond warm glow utility and compensation. Research participation can serve as the means for satisfying many of the "basic human goods" in which some theorists say welfare inheres. According to John Finnis's natural law theory, the "basic human goods" are life, knowledge, play, aesthetic experience, friendship or sociality, practical reasonableness, and religion.[143] Martha Nussbaum's "capabilities approach" proposes a similar list of "human capabilities" or "substantive freedoms" central to human flourishing: life; bodily health; bodily integrity; senses, imagination, and thought; emotions; practical reason; affiliation; other species; play; and control over one's environment.[144] And hundreds of motivational studies conducted under the rubric of self-determination theory[145] suggest that human beings are intrinsically motivated by their drive to satisfy three basic psychological needs: competence,[146] relatedness,[147] and autonomy and

142. M.J. Czarny et al., *Assessing Payment to Healthy Volunteers in Clinical Research: The Research Subject's Perspective*, 87 CLINICAL PHARMACOLOGY & THERAPEUTICS 286 (2010).

143. *See* JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS 90 (1980).

144. Martha C. Nussbaum, *Capabilities and Human Rights*, 66 FORDHAM L. REV. 273, 287–88 (1997).

145. A theory under which the individual neglects the other half emphasized by existentialists—namely, that we have psychological needs to experience ourselves not only as autonomous but also, in different circumstances, as determined by forces outside our control (thus avoiding the unpleasant burden of responsibility).

146. In this context, competence refers to the need to experience oneself as capable and competent in controlling the environment and being able to reliably predict outcomes and belief that you can do something well. Edward L. Deci & Maarten Vansteenkiste, *Self-Determination Theory and Basic Need Satisfaction: Understanding Human Development in Positive Psychology*, 27 RICERCHE DI PSICOLOGIA 23, 25 (2004).

147. Relatedness is the need to believe that you matter and that others matter to you, to care for and be related to others, including the need to experience authentic relatedness from others and to experience satisfaction in participation and involvement with the social world. *See* Roy F. Baumeister & Mark R. Leary, *The Need to Belong: Desire for Interpersonal Attachments as a Fundamental Human Motivation*, 117 PSYCHOL. BULL. 497 (1995).

self-determination.[148]

Research participation can serve several of these ends. For instance, participants may receive many of the same psychic benefits that drive researchers to conduct research, such as the satisfaction of intellectual curiosity and of gaining a sense of control through prediction. Women who had recently experienced intimate partner abuse were asked to list the reasons they had decided to participate in a longitudinal study of such abuse. The most common reason—cited by 66.9% of respondents as one of the top three reasons they enrolled—was curiosity.[149]

As for control, recall the earlier discussion of research on sensitive topics, where IRBs (accurately) worried particularly about the risk of retraumatizing participants by asking them to recall painful pasts.[150] Researchers have found that, although participants who have experienced painful pasts—for instance, child abuse—are more likely (compared to other participants) to report distress when asked by researchers to remember those pasts, these participants are also more likely to report that participation *benefitted* them.[151] On reflection, this is not all that surprising. Many people find it therapeutic to discuss painful experiences. A participant's particular painful experience may be viewed skeptically by laypersons or even mainstream experts, and so they may feel validated by professional research interest in those experiences. Or they may have experienced something, such as an assault, that left them feeling disempowered, and so contributing to attempts to better understand the causes and effects of that assault may provide them with a sense of empowerment.

C. Reasonableness of Risk–Benefit Tradeoffs

The final step in research risk–benefit analysis is determining whether a research project's "[r]isks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result."[152] What does it

148. Autonomy is the need to actively participate in determining one's own behavior, including the need to experience one's actions as the result of autonomous choice without external interference, the need to believe that you have a say in how you live. *See* Edward L. Deci & Richard M. Ryan, *Human Autonomy: The Basis for True Self-Esteem*, *in* EFFICACY, AGENCY, AND SELF-ESTEEM 31–49 (Michael H. Kernis ed., 1995).

149. Claire L. Hebenstreit & Anne P. DePrince, *Perceptions of Participating in Longitudinal Trauma Research Among Women Exposed to Intimate Partner Abuse*, J. EMPIRICAL RES. ON HUM. RES. ETHICS, April 2012, at 60, 64.

150. *See supra* Part II.A.1.a.

151. Decker et al., *supra* note 75, at 62.

152. 45 C.F.R. § 46.111(a)(2) (2012).

mean to find that a study's risk–benefit ratio is "reasonable"?

1. What Risk–Benefit "Reasonableness" in HSR Is Not

Many have rightly noted the vagueness of this federal research risk–benefit reasonableness standard,[153] which is more or less the same in the regulatory frameworks of other countries and in international codes of research ethics.[154] The regulatory standard tells IRBs which components to include in its risk–benefit calculus—risks to participants, expected benefits to participants, and the importance of the knowledge reasonably expected to result from the research—but not the relative weights that IRBs should assign to each of these components. Still, we can eliminate two seemingly plausible reasonableness standards.

a. Reasonableness as Social Welfare Maximization

One possibility is that HSR risk–benefit analysis should be understood to refer to the standard cost-benefit analysis (CBA) that is ubiquitous in the regulatory state. Indeed, some scholars, especially those from the "overregulation camp," assume that maximizing overall social welfare is the goal of the current system of research governance, or should be, without grappling with the fairly radical normative shift this would entail.[155]

---

153. *See* Coleman, *supra* note 12, at 14 ("Each [IRB] member is free to interpret this reasonableness standard as he or she sees fit."); *id.* at 20 (arguing that the "reasonable risk" standard is "inherently amorphous" and "susceptible to a virtually limitless range of possible interpretations").

154. *See infra* text accompanying notes 169–71.

155. *See, e.g.*, Carl H. Coleman, *Vulnerability as a Regulatory Category in Human Subject Research*, 37 J.L. MED. & ETHICS 12, 15 (2009) (according to "the basic 'deal' that underlies society's regulation of human subject research," research need not be in a participant's best interests, but a study's "objective risk–benefit ratio" must yield a "net social benefit"); *id.* at 16 ("[Risks] need not be an absolute barrier to proceeding with research, as long as the expected benefits of the study outweigh the study's overall risks."); *see also* Ernest D. Prentice & Dean L. Antonson, *A Protocol Review Guide To Reduce IRB Inconsistency*, IRB: ETHICS & HUM. RES., Jan.–Feb. 1987, at 9 (IRBs should ask whether risks are "balanced or outweighed" by benefits); Zywicki, *supra* note 15, at 865 (assuming, "[f]or purposes of this Article," that the regulatory goal "is to maximize the net social benefits of the governance of academic research by minimizing the costs of Type I and Type II errors and administrative costs"); *id.* at 883 ("Like governmental safety regulations, the objective function of IRBs is to minimize the costs of the IRB system through the minimization of Type I and Type II errors as well as administrative costs."); Hyman, *supra* note 15, at 753 ("The goal for any system of research oversight is to maximize the number of true positives and negatives . . . , and minimize the number of false positives and false negatives . . . , and the costs of research oversight. These costs include the transaction costs of operating the system . . . , and the costs of erroneous decisions and delay."); Hamburger, *supra* note 14, at 469 (arguing that "the loss to humanity over the decades" from IRB alterations of research "almost certainly exceeds the loss from

Let us assume, *arguendo*, that IRBs are capable of developing a risk–benefit profile for each study that accurately reflects the expected costs and benefits of participation for all prospective participants. Interpreting risk–benefit reasonableness as CBA would mean that research is "reasonable" whenever its expected benefits to society and to participants, "if any," *outweigh* its risks to participants.[156] This regulatory construction would limit the heterogeneity problem to earlier stages of risk–benefit analysis. IRBs would still suffer from an information problem at the level of risk and benefit assessment, since the social welfare function is merely the aggregation of the individual welfare functions of all members of society. But, having aggregated risks to participants on one side and expected benefits to them and society on the other, "reasonableness" would become a simple matter of math: the regulations, so constructed, would direct the IRB to choose whichever alternative—approving or disapproving the study—maximizes net social welfare, regardless of[157] the distribution of the costs and benefits of that choice.

But risk–benefit analysis in HSR is not, and is not intended to be, standard CBA. For one thing, HSR risk–benefit analysis is not a plenary assessment of the expected costs and benefits of research.[158] More importantly, the governance of HSR is decidedly non-utilitarian. The primary problem with the research abuses that prompted the National Research Act was not their disproportionate distribution of risks and expected benefits. Research by definition[159] entails participants assuming

---

the research by a very substantial factor").

156. Rather than a purely utilitarian governance scheme marked by conscription into research, social welfare maximization through research might be subject to an (almost) absolute side constraint of voluntary, informed consent.

157. Or at least separate from this distribution; in theory, welfare economics is not indifferent to the distribution of costs and benefits, but sees merit in separating the economic task of maximizing surplus value from the political task of redistributing the resulting wealth. *See* Adler & Posner, *supra* note 138, at 185–86.

158. Standard CBA considers the welfare of all individuals affected by a decision and seeks to maximize net aggregate welfare. HSR regulations direct IRBs to consider more or less immediate risks to participants, benefits to participants, and benefits to society in the form of knowledge production. But IRBs are implicitly or explicitly directed to ignore costs to any party other than participants, long-term risks to participants from the potential policy implications of the knowledge gained through research, and benefits to nonparticipants such as academic freedom or employment for researchers. Even as to the category of expected benefits to participants, IRBs are directed not to count the vast majority of benefits to participants. *See infra* Part III.B.

159. *See* BELMONT REPORT, *supra* note 24, at 2–3 ("'[P]ractice' refers to interventions that are designed solely to enhance the well-being of an individual patient or client and that have a reasonable expectation of success. . . . By contrast, the term 'research' designates an activity designed to test an hypothesis, permit conclusions to be drawn, and thereby to

the risks of activities whose resulting benefits in the form of the public good of knowledge will necessarily be largely (in the case of therapeutic research) or completely (in the case of nontherapeutic research) externalized on others. Rather, the fundamental problem with these studies was that they *conscripted unwitting* individuals into research, thus making the classic utilitarian mistake of subsuming their individual preferences to the pursuit of socially useful knowledge.[160] Individuals can and do assume risks despite the fact that they themselves are not expected to benefit from that activity, or that they are expected to benefit far less than others. The voluntary, informed consent of that individual transforms a utilitarian decision imposed upon them into one that continues to be expected to maximize social welfare, but does so through a process in which the risk-assuming individual adopts that end as her own.[161]

Finally, if the regulations intended for any research that maximizes social welfare to be approved, IRBs risk–benefit analysis would hardly be necessary. The aggregation across all members of society of even very modest welfare gains from the production of useful knowledge will easily outweigh even devastating setbacks to the welfare of a relative few participants, especially if one counts the benefits of knowledge production to future generations. Although much research produces *de minimus* social value, much more produces modest benefits and some fraction produces almost incalculable social benefits. Because it is difficult, if not impossible, to determine *ex ante* which research will be wildly beneficial, which will fail

---

develop or contribute to generalizable knowledge . . . ."); 45 C.F.R. § 46.102(d) (2012) (defining *research* as activities that are, *inter alia*, "designed to develop or contribute to generalizable knowledge"); *id.* § 46.111(a)(2) (implying a narrow understanding of expected benefits to participants, for risk–benefit analysis purposes, by qualifying it with "if any"). The current system of research governance was premised on an explicit distinction between research and professional practice. According to this distinction, professional practice involves the single-minded pursuit of the individual patient's (or client's) best interests. Research, by contrast, aims to produce generalizable knowledge that is primarily of value to society and future patients (or clients); any direct benefit to subjects is merely a happy coincidence.

160. In "adopt[ing] for society as a whole the principle of rational choice for one man," utilitarianism fails to "take seriously the distinction between persons." JOHN RAWLS, A THEORY OF JUSTICE 26–27 (1971) (drawing on Kant's notion of respect for persons); *see* Alexander Volokh, Commentary, *Rationality or Rationalism? The Positive and Normative Flaws of Cost–Benefit Analysis*, 48 HOUS. L. REV. 79, 90 (2011) (noting "how seamlessly one slips from the *we* who pay and receive to the *we* who receive and forego. But these are different groups, and there is neither a common pocketbook nor a common valuing mind").

161. "Act in such a way that you treat humanity, whether in your own person or in the person of another, always at the same time as an end and never simply as a means." IMMANUEL KANT, GROUNDING FOR THE METAPHYSICS OF MORALS 36 (James W. Ellington trans., 3d ed. 1993) (1785) (Second Formulation of the categorical imperative).

entirely, and which will fall somewhere in between, the anticipated benefits of research must be viewed in the aggregate, across all protocols (at least within broad categories of research).

b. Reasonableness as Participant Welfare Maximization

At the other extreme, we could imagine a world in which a decision to participate in research is "reasonable"—that is, neoclassically "rational"—only when enrollment is expected to be in the participant's "best interests"—that is, to maximize her welfare relative to her alternatives. Indeed, some scholars, especially those from the "underregulation camp,"[162] some international codes of research ethics,[163] and much of the conventional wisdom of research ethics,[164] imply that the participant–

162. *See, e.g.*, Tsiao Yi Yap et al., *Both Sides of the Coin: Randomization from the Perspectives of Physician-Investigators and Patient-Subjects,* 20 ETHICS & BEHAV. 380, 384 (2010) ("By reassuring patients explicitly and directly that their own well-being will *always* come before the scientific goals of the research, physician investigators can build trust in the context of the [randomized controlled trial (RCT)].").

163. *See, e.g.*, INT'L CONFERENCE ON HARMONIZATION OF TECHNICAL REQUIREMENTS FOR REGISTRATION OF PHARM. FOR HUMAN USE, ICH HARMONIZED TRIPARTITE GUIDELINE: GUIDELINE FOR GOOD CLINICAL PRACTICE E6(R1), princ. 2.3 (1996) [hereinafter ICH] ("The rights, safety, and well-being of the trial subjects are the most important considerations and should prevail over interests of science and society.").

164. For instance, it has long been orthodoxy in research ethics that RCTs are acceptable only if the relevant expert community is in "equipoise" regarding both (1) the value to the participant of being randomized into the treatment arm versus the control arm of the study (internal equipoise); and (2) the value to the participant of being enrolled in either of these research arms versus the value of the alternatives available to her outside of the study (external equipoise). *See* CHARLES FRIED, MEDICAL EXPERIMENTATION: PERSONAL INTEGRITY AND SOCIAL POLICY 51 (1974) (recognizing that the RCT presents the physician–researcher with a conflict of interest, and coining the term "equipoise" to refer to situations when the investigator has no professional reason to prefer one treatment to another); Benjamin Freedman, *Equipoise and the Ethics of Clinical Research*, 317 NEW ENG. J. MED. 141, 141 (1987) (coining "clinical equipoise" as the requirement that the "expert medical community," rather than the individual professional, be indifferent between treatments and that the absence of clinical equipoise renders an RCT unethical). Thus, research participation need not be the superior option for participants, but it must not be, to the best knowledge of experts, an inferior option. As some scholars have begun to note, such a standard is difficult to square with the *Belmont Report*'s view of participants as volunteers, and converts research into quasi-therapy. *See, e.g.*, Steven Joffe & Franklin G. Miller, *Bench to Bedside: Mapping the Moral Terrain of Clinical Research*, HASTINGS CTR. REP., Mar.–Apr. 2008, at 30, 31 ("[R]esearch ethics is characterized by a basic incoherence: on one hand, clinical research is seen as ethically distinct from medical care; on the other hand, the obligations of investigators, especially in clinical trials, are thought to be grounded in the ethics of the doctor-patient relationship."); *see also* Franklin G. Miller & Howard Brody, *A Critique of Clinical Equipoise: Therapeutic Misconception in the Ethics of Clinical Trials*, HASTINGS CTR. REP., May–June 2003, at 19. In the context of research in developing countries, similar proposals

researcher relationship is fiduciary.[165] As such, reasonable research is that which meets a "best-interests-of-the-participant" standard, where a study's expected benefits to participants match or exceed its risks to them, or where its risk–benefit profile is at least as favorable as participants' alternatives.

But if this were the criterion of research reasonableness, then—at least on the current account of research-related benefits, which essentially recognizes only direct clinical benefit to participants as benefits[166]—only relatively rare "therapeutic" research (often involving experimental interventions to address conditions for which no standard treatment exists) would be reasonable. This standard would preclude the altruism that both the *Belmont Report* and the regulations strongly associate with research participation.[167]

This standard of participant welfare maximization, under which virtually no research is reasonable, is thus just as implausible an interpretation of the regulations' requirement of risk–benefit "reasonableness" as is a social welfare maximization standard, under which virtually *all* research is reasonable. Both make IRB risk–benefit analysis superfluous.

2. Heterogeneity in Risk–Benefit Tradeoff Preferences

The Commission may have recognized the implausibility of both of the above criteria for reasonableness, under which research is acceptable only if its expected benefits (whether for participants and society, or solely for participants) outweigh its risks. HHS's earlier regulations had in fact required IRBs to find that "[t]he risks to the subject are *so outweighed* by the sum of the benefit to the subject and the importance of the knowledge to be gained as to warrant a decision to allow the subject to accept these risks."[168] Various codes of the 1960s and 1970s similarly required that research benefits outweigh risks.[169] But the Commission recommended that this rule

include those that would require researchers to provide participants with medical care that is ancillary to the trial, or with post-trial access to an experimental drug.

165. For rejection of the fiduciary view of the participant–researcher relationship, see Michelle N. Meyer, *The Subject–Researcher Relationship: In Defense of Contracting Around Default Rules*, AM. J. BIOETHICS, Apr. 2011, at 27.

166. *See infra* Part III.B.

167. *See supra* text accompanying note 156.

168. 45 C.F.R. § 46.102(b)(1) (1978) (emphasis added).

169. *See* BELMONT REPORT, *supra* note 24, at 16. For instance, the Declaration of Helsinki provides that medical research should be conducted only if "the importance of the objective outweighs the inherent risks and burdens to the research subjects," and halted "if the risks are found to outweigh the potential benefits." World Med. Ass'n, *Declaration of Helsinki: Ethical Principles for Research Involving Human Subjects*, ¶¶ 20, 21 (Oct. 2008). Principle 6 of the Nuremberg Code of 1947 similarly requires social benefits to be equal to or greater than risks to participants: "The degree of risk to be taken should never exceed that

be replaced with the current "reasonableness" standard.[170] Today, other regulatory and ethical guidelines around the globe similarly call for risk–benefit ratios to be "proportionate,"[171] "favorable,"[172] or "justified."[173] These standards set no cap on the amount of risk that (nonvulnerable) participants are allowed to assume. In theory, at least, very large risks could be "reasonable in relation to" expected benefits so long as a study has great social promise.

Such standards—when employed by central actors such as IRBs—falter on the shores of participant heterogeneity. To say that something is "reasonable" entails a claim that it is reasonable *to* someone; like the proverbial tree that falls silently in the forest, things are not reasonable in the abstract. The question is: *To whom* must they be reasonable? Given the analysis in Section C.1 above, the reasonableness of research risk–benefit profiles presumably should be determined not from society's perspective, but from the individual participant's perspective. But, except for extreme cases of disproportionate risks and benefits (which competent prospective participants are extremely unlikely to accept), individuals will reasonably disagree about the "reasonableness" of assuming some quantity and kind of risk in pursuit of some quantity and kind of benefit for themselves or others. Let us stipulate that a study offers two prospective participants an identical risk–benefit profile. We can easily imagine that one finds this distribution of risks and benefits to be reasonable while the other does not. Prospective participants will almost certainly be heterogeneous with respect to how much benefit must be expected, and in what proportion of benefits to themselves versus benefits to society, before the risks on the other side of the

---

determined by the humanitarian importance of the problem to be solved by the experiment." 2 TRIALS OF WAR CRIMINALS BEFORE THE NUREMBERG MILITARY TRIBUNALS UNDER CONTROL COUNCIL LAW NO. 10, 181–82 (U.S. Gov't Printing Off. 1946–1949). Principle 10, however, caps acceptable risk by requiring termination of a study that it is "likely to result in injury, disability, or death" to the participant. *Id.* at 182.

170. *See* NAT'L COMM'N, IRBS, *supra* note 254, at 20; 45 C.F.R. § 46.111(a)(2) (2012); *see also* COUNCIL FOR INT'L ORGS. OF MED. SCIS. (CIOMS), INTERNATIONAL ETHICAL GUIDELINES FOR BIOMEDICAL RESEARCH INVOLVING HUMAN SUBJECTS, 47 (3d ed. 2002) [hereinafter CIOMS] (risk–benefit ratio should be "reasonable").

171. World Med. Ass'n, *Declaration of Helsinki, Recommendations Guiding Physicians In Biomedical Research Involving Human Subjects*, princ. 4 (Sept. 1989); Convention for the Protection of Human Rights and Dignity of the Human Being with Regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, art. 16, *opened for signature* Apr. 4, 1997, E.T.S. No. 164 (entered into force Dec. 1, 1999) (ratified by twenty-nine countries).

172. *See* ROBERT J. LEVINE, ETHICS AND REGULATION OF CLINICAL RESEARCH 63 (2d ed. 1986); MED. RESEARCH COUNCIL OF CAN., GUIDELINES FOR RESEARCH INVOLVING HUMAN SUBJECTS 12 (1987).

173. ICH, *supra* note 1633, princ. 2.2.

ledger become "reasonable" to them.[174] This is because even if risk assessment is the province of fact, risk *acceptability* is the province of value.[175] But IRBs lack access to such private information and, in any case, would be unable to make more than a single reasonableness determination that would be binding on all prospective participants.

Empirical research, not surprisingly, bears this out. For example, one study found that cancer patients were willing to undergo chemotherapy for a mere 1% chance of cure, while doctors would need a 10% chance, and both oncology nurses and members of the general public would require a 50% chance.[176] Notice that these differences cannot be explained by whether the respondents had been exposed to the effects of chemotherapy; both those who required the greatest probability of benefit—cancer nurses—and those who required the least—cancer patients—have almost certainly had considerable exposure to the side effects of chemotherapy. Moreover, cancer nurses and members of the general public—groups that presumably have, on average, very different levels of exposure to chemotherapy's effects—gave identical answers.

In a Canadian study, forty-four biomedical IRBs reviewed a hypothetical protocol in which participants would be studied using fMRI in order to identify the neurobiological correlates of social behavior.[177] Participants would be scanned, and their brain activity observed, while they were in three "states": an ordinary state, in which participants would be asked to think about everyday things such as grocery lists; a meditation state, in which they would be asked to introspect; and a violent state, in which they would view photographs showing "extreme violence." Of the forty-four IRBs, three (7%) would have approved the study unconditionally, ten (22.7%) would have done so conditionally (requiring alterations of greater or lesser significance to the protocol), twenty-three (52%) would have given the study a qualified rejection, and seven (16%) would have given it an unqualified rejection.[178] As the authors of the study observe, the "concerns

174. Heterogeneity in other meta-preferences is also likely to be relevant to research participation. For instance, individuals have different discount rates: even assuming that everyone equally values immediate outcome X, they will often differ as to how much they discount that value when X is not immediate but is more or less delayed. *See* GEORGE AINSLIE, PICOECONOMICS: THE STRATEGIC INTERACTION OF SUCCESSIVE MOTIVATIONAL STATES WITHIN THE PERSON 363–64 (1992).

175. *See, e.g.*, NAT'L RES. COUNCIL, TECHNICAL BASES FOR YUCCA MOUNTAIN STANDARDS 5 (1995) ("[D]etermining what risk level is acceptable is not ultimately a question of science but of public policy.").

176. Agrawal & Emanuel, *supra* note 18, at 1077–78.

177. J. de Champlain & J. Patenaude, *Review of a Mock Research Protocol in Functional Neuroimaging by Canadian Research Ethics Boards*, 32 J. INST. MED. ETHICS 530 (2006).

178. *Id.* at 532 tbl. 5.

raised by 23 of the 30 rejecting [IRBs] resembled, overall, the concerns mentioned by the 10 [IRBs] that handed down conditional approvals. Thus, certain [IRBs] shared a set of concerns; yet, some approved the project and others rejected it."[179] A more plausible explanation for this outcome than IRB incompetence is variation among IRB members in risk–benefit tradeoff preferences.

In a Netherlands study, forty-three IRB members were asked to assess the risks and expected benefits of a hypothetical breast cancer study. Thirty percent said that the risks outweighed the expected benefits. Twenty-one percent said that the expected benefits outweighed the risks. Thirty-five percent said the risks and expected benefits were about equal.[180]

III. IRB RESPONSES TO THE HETEROGENEITY PROBLEM AND THEIR COSTS: THE EGGSHELL PARTICIPANT

Although IRBs, regulators, and research ethicists rarely, if ever, acknowledge the heterogeneity problem, they have developed ways of assessing research risks and expected benefits that, I argue, are implicit responses to participant heterogeneity. This Part details these responses and their costs.

A. Responses to Risk Heterogeneity

IRBs are directed to consider risks of virtually every nature, including (but not limited to) those that are physical, psychological, social, economic and legal,[181] posed by any aspect of the research protocol.[182] They consider risks of *de minimis* harm, including *discomfort*—"unpleasant sensations or emotions of short duration and minimum to moderate severity," such as "shortness of breath in a study of maximal exercise tolerance or the anxiety associated with being stuck with a needle";[183] *inconvenience*—"any interference with the subject's ability to carry out usual activities," such as

179. *Id.* at 533. One board (2.3%) did not answer. *Id.* at 532 tbl. 5.

180. van Luijn et al., *supra* note 9, at 170, 172.

181. *See, e.g.*, N.Y. STATE DEP'T OF HEALTH, SAFEGUARDING HEALTHY RESEARCH SUBJECTS: PROTECTING VOLUNTEERS FROM HARM 7 (1999) [hereinafter N.Y. IRB Guidelines] ("[R]isk refers to any physical, psychological, economic, social, *or other* harm associated with the research. In assessing risk, the IRB should consider . . . *any* harm that subjects might experience . . . ." (emphases added) (footnote omitted)). The only risks to participants that IRBs are directed to ignore are the "possible long-range effects of applying knowledge gained in the research (for example, the possible effects of the research on public policy)." 45 C.F.R. § 46.111(a)(2) (2012).

182. *See* N.Y. IRB Guidelines, *supra* note 181, at 8 ("All elements of the protocol should be subjected to risk analysis.").

183. *Id.* at 9.

"time spent travelling to the research facility or the need to get up in the middle of the night to take a pill";[184] and mere "boredom."[185] As such, there is effectively no such thing as no-risk research.

IRBs consider any conceivable risk faced by any conceivable prospective participant. IRBs rarely rely on data regarding the population frequency of these harms, but instead tend to proceed as if the risk faced by this "eggshell" participant were the risk faced by all (or even the modal) prospective participants. IRBs tend to speculate about the likelihood and magnitude of research-related harms and, in many cases, the data suggest that the risks IRBs take very seriously are visited upon relatively few research participants.

B. Responses to Benefit Heterogeneity

IRBs deal very differently with participant heterogeneity in the context of expected benefits. The *Belmont Report* admonishes that "[m]any kinds of possible harms *and benefits* need to be taken into account. There are, for example, risks of psychological harm, physical harm, legal harm, social harm and economic harm *and the corresponding benefits*."[186] The *IRB Guidebook*, moreover, reasonably defines "benefit" as any "valued or desired outcome; an advantage."[187] Yet other agency guidance, as well as the canonical norms of research ethics, direct IRBs not to count the vast majority of what in fact motivates participants to enroll in research, including compensation (both monetary and in-kind), altruism (both "pure," disinterested altruism and warm glow utility), other forms of psychic income (e.g., curiosity, empowerment), and even medical benefits that are ancillary to the research (e.g., diagnostics, examinations, consults). Essentially, IRBs systematically count only expected medical benefits to participants in their risk–benefit analysis (which are obviously not a possibility in the vast majority of regulated HSR that is not biomedical in

184. *Id.*

185. *See also* Haggerty, *supra* note 37, at 400 ("The range of potential research related harms envisioned by [Canadian IRBs] at times seems to be limited only by the imagination of different reviewers. Any change in a research participant's condition or disruption of their routine can be conceived of as a potential harm . . . [including] such things as the possibility that a research participant might be embarrassed by personal questions or that they might experience disruption in their family routine or a loss of respect by others.").

186. BELMONT REPORT, *supra* note 24, at 15 (emphases added).

187. OFFICE FOR HUMAN RESOURCE PROTECTION (OHRP), IRB GUIDEBOOK ch. III.A (1993) [hereinafter IRB GUIDEBOOK]; *see also* BELMONT REPORT, *supra* note 24, at 15 ("The term 'benefit' is used in the research context to refer to something of positive value related to health or welfare.").

nature) and the system tends to undercount even these.[188] This impoverished view of human benefit is consistent with the neglect of research benefits relative to research risks.[189]

New York State's guidelines for conducting HSR with healthy volunteers makes particularly explicit the conventional wisdom:

> [In HSR,] 'benefit' . . . refers to two distinct factors: (1) any direct enhancement to the health and well-being of the individual subject (not at issue in [research involving healthy volunteers]); and (2) the prospect of increasing knowledge of benefit to society. For research not designed to provide a direct benefit to the individual subjects, the only relevant benefit is the prospect of increasing knowledge.
>
> IRBs should recognize that there are certain aspects of research that subjects are likely to perceive as benefits but that do not fall within the definition of 'benefit' set forth above. For example, subjects may derive altruistic satisfaction from the expectation that others, including family and friends, may benefit from the knowledge that might be gained from the study. In addition, subjects may benefit financially from participating in the research. While some subjects may attach considerable importance to these factors, they do not constitute the type of 'benefit' that IRBs should consider in evaluating the risk–benefit ratio of a protocol.[190]

1. Altruism and Other Intangible Benefits

Many research participants cite the desire to help others as among the reasons they participate in research. The concept of altruism has, moreover, played a substantial role in the development of research governance. The *Belmont Report* characterizes each research participant as, "in essence a volunteer."[191] Research ethicists, similarly, have tended to insist that participants should be motivated by altruism, rather than by either compensation or anticipation of direct therapeutic benefit.

Yet, like many other social systems, our research governance system is

---

188. *See infra* subpart III.B.3.

189. *See* Nancy M.P. King, *Defining and Describing Benefit Appropriately in Clinical Trials*, 28 J.L. MED. & ETHICS 332, 332 (2000) ("Everyone on an IRB has probably spent time . . . arguing over whether a three-page bulleted list of risk description is helpful or overkill . . . . Yet only a small fraction of [that] time . . . has been devoted to discussing whether and when potential benefit to subjects can reasonably be claimed and, if so, how it should be described . . . ."); Larry R. Churchill et al., *Assessing Benefits in Clinical Research: Why Diversity in Benefit Assessment Can Be Risky*, IRB: ETHICS & HUM. RES., May–June 2003, at 1, 1 ("Despite its importance, this topic has received little attention in the bioethics literature. . . . [A]ttention to risks typically overshadows discussion of benefits.").

190. N.Y. IRB Guidelines, *supra* note 181, at 10 (footnote omitted).

191. *See* BELMONT REPORT, *supra* note 24, at 11.

poorly designed to accommodate pro-sociality.[192] In their risk–benefit analysis, IRBs do not count participants' preferences for altruism. A policy not to count "pure," other-regarding altruism might otherwise be explained by a desire to avoid "double-counting" research benefits to society.[193] But IRBs also decline to count the warm glow utility research participants often receive. Nor do IRBs count other intangible benefits to participants, such as satisfaction of intellectual curiosity and increased sense of self-respect or empowerment.

2. Payment and Compensation In-Kind

In other employment contexts, when individuals agree to perform risky tasks in order to benefit others (namely, their employer), the terms of their employment contract will, barring market failure, reflect some sort of "compensating differentials" (wage premiums or nonmonetary benefits). In the U.S., this hazard pay costs employers an estimated $245 billion each year (in 2004 dollars)—over 2% of the gross domestic product and 5% of total wages.[194] Yet any payment that research participants receive does not count as a benefit to them in IRB risk–benefit analysis. According to OHRP's *IRB Guidebook*, "Direct payments or other forms of remuneration offered to potential subjects as an incentive or reward for participation should not be considered a 'benefit' to be gained from research."[195]

Indeed, IRBs often view payment as an undue inducement,[196] or coercive,[197] and thus as a risk, of sorts. For instance, many IRBs and

192. *See generally* BENKLER, *supra* note 129 (arguing that the "myth" of unbounded self-interest has shaped the design of social systems, which should be redesigned according to the reality of pro-sociality, as demonstrated by research in neuroscience, evolutionary biology, business and engineering that finds substantial cooperation among humans).

193. *Cf.* MATTHEW D. ADLER & ERIC A. POSNER, NEW FOUNDATIONS OF COST-BENEFIT ANALYSIS 133–36 (2006) (arguing that all "disinterested preferences," including altruism, should be excluded from cost–benefit analysis).

194. And that does not include additional "risk costs" to employers in the form of nonmonetary benefits, or higher worker's compensation premiums. *See generally* W. Kip Viscusi, *Job Safety*, *in* The CONCISE ENCYCLOPEDIA OF ECONOMICS (2d ed. 2008), *available at* http://www.econlib.org/library/Enc/JobSafety.html.

195. IRB Guidebook, *supra* note 1877 at ch. III.A; *see also* FDA, INFORMATION SHEET GUIDANCE FOR INSTITUTIONAL REVIEW BOARDS (IRBS), CLINICAL INVESTIGATORS, AND SPONSORS: PAYMENT TO RESEARCH SUBJECTS (1998), *available at* http://www.fda.gov/RegulatoryInformation/Guidances/ucm126429.htm.

196. *See, e.g.*, CIOMS, *supra* note 170, at 45; BELMONT REPORT, *supra* note 24, at 10–14.

197. For example, one IRB determined that paying undergraduates $25 per quarter to fill out interview forms would be "coercive" for financially squeezed students who would not feel free to refuse the offer. "The IRB insisted that the payment be reduced to $10 a quarter, thus protecting a bunch of students from making a little bit of pocket money while destroying the utility of the survey by causing a sufficiently high non-response rate to cast

research ethicists deem the healthy volunteers who participate in first-in-human and other Phase I drug trials to be vulnerable to undue inducement because they are presumably motivated solely by payment.

Payment in kind, too, is not counted as a benefit to participants and is often considered borderline coercive. For instance, IRBs frequently express concern that offering students an opportunity to earn extra credit by participating in research would constitute the proverbial offer they can't refuse.[198]

3. Medical Benefits

The one category of benefits to participants that IRBs *are* often willing to "count," at least in theory, is medical benefits. But the extent to which IRBs count even these benefits is significantly limited by at least two factors. First, research regulators and ethicists tend to downplay the extent to which research carries the prospect for medically benefitting participants. And so, even when medical benefits technically count, they are often underweighted, sometimes to the point of rating a "zero" on the utility scale. The notion that research is not designed to benefit participants is built into the very regulatory definition of "research." Research ethicists and regulators have long worried that participants in medical research suffer from the "therapeutic misconception," in which they confuse research and the practice of medicine, and as a result "deny the possibility that there may be major disadvantages to participating in clinical research that stem from the nature of the research process itself."[199] To offset this possibility, IRBs downplay the possible clinical benefits of biomedical research to prospective participants, and even seem to internalize this thinking themselves.

For instance, IRBs and research ethicists often claim that there is "no reasonable probability" that participation in a Phase I oncology trial will benefit cancer patients. This, combined with a 0.5% risk of toxicity-induced death, sometimes leads commentators to conclude that the risk–benefit profile of such trials is so unfavorable that the trial "may not be

---

doubt on whether the respondents were representative of the original sample. All this despite the fact that the proposed study *posed no risk whatever to its subjects*." Mark Kleiman, *Bleg: IRB Horror Stories*, THE REALITY-BASED COMMUNITY BLOG, (Apr. 14, 2009, 7:14 PM) http://www.samefacts.com/2009/04/science-and-its-methods/bleg-irb-horror-stories/.

198. *See* IRB GUIDEBOOK, *supra* note 187, at ch. VI.J (noting that "[r]equiring participation in research for course credit (or extra credit) is . . . controversial" and suggesting "mechanisms . . . for diminishing or eliminating the coercive aspect of student participation for course credit that IRBs might find useful").

199. Paul S. Appelbaum et al., *False Hopes and Best Data: Consent to Research and the Therapeutic Misconception*, HASTINGS CENTER REP., Apr. 1987, at 20, 20.

performed."[200] If they consider benefits at all, IRBs are likely to turn to oft-cited meta-analyses of Phase 1 oncology studies that show that 5% of patient-participants respond to the investigational drug. But "such aggregate data conceal important information": More than 60% of the compounds tested had at least one "objective response" (i.e., tumor shrinkage of more than 50%), more than 30% of experimental drugs had greater than a 5% response rate and, in some cases, "substantial clinical benefits and even cures have been achieved."[201] For example, in the 1970s, cisplatin for testicular cancer "had a response rate of more than 50%, and in a quarter of cases the tumor completely disappeared and was probably cured."[202] And, more recently, imatinib mesylate for chronic myeloid leukemia "demonstrated complete hematologic response rates of 98%, of which 96% lasted beyond 1 year."[203]

Second, research regulators and ethicists tend to count only "direct," and not "indirect," benefits of all kinds, including medical benefits. Direct benefits are those that derive from the research procedures themselves. Indirect benefits, by contrast, are ancillary to a study, and include not only payment and compensation in-kind but also medical examinations, medicines, and psychological counseling. Thus, in a survey of IRB chairpersons, only 51% thought that added medical examinations and medicines offered the prospect of direct benefit to participants, and only 60% thought that added psychological counseling did so. These already relatively low percentages flout the regulations' definition of "direct benefit," OHRP's interpretation of the regulations in its *IRB Guidebook*, and the view of "most commentators."[204]

C. IRB Risk-Aversion and Its Costs

One possible implicit rationale for the twin practices of overemphasizing risks and undercounting expected benefits to participants is that participant heterogeneity, as we have seen, makes it inappropriate for IRBs to assume that all participants would receive the same (or any) benefit. But this "solves" the heterogeneity problem by assigning *no* utility to, say, compensation or altruism for *any* participant and by assigning to *all* prospective participants the disutility of *every* conceivable cost to *any*

200. Agrawal & Emanuel, *supra* note 18, at 1076 (quoting George J. Annas, *The Changing Landscape of Human Experimentation: Nuremberg, Helsinki, and Beyond*, 2 HEALTH MATRIX 119 (1992)).

201. *Id.*

202. *Id.*

203. *Id.*

204. *See* Shah et al., *supra* note 76, at 478, 480.

conceivable participant. The net effect is that IRB risk–benefit decisionmaking is systematically risk-averse relative to the utility that at least some, and very often most, prospective participants could expect from enrolling in a study.

The costs of this risk aversion are hardly trivial. Although IRBs do not often reject research proposals outright, the most comprehensive data available[205] suggest that IRBs require alterations to over 80% of the submissions they eventually approve—about 100,000 altered proposals each year.[206] The data do not indicate the nature of these alterations, and most are likely designed to enhance the informed consent process or minimize gratuitous risks, rather than to alter the risk–benefit profile of the study.[207] Yet IRBs do sometimes require substantive alterations that render research less valuable, more costly, or more difficult (or practically impossible) to conduct.[208]

---

205. BELL ET AL., *supra* note 45, at 4. The *Bell Report*, a National Institutes of Health-commissioned survey of the 1995 practices of 491 IRBs, found that 94% of IRBs were equally or more likely to require changes to a proposal than to approve it as submitted. *Id.* at 11 fig. 4. Thirty-four percent did not approve any proposals without alterations that year. *Id.* at 29 fig. 15. Overall, 73% of IRBs approved 25% or fewer protocols as submitted; 34% approved none as submitted; 10% approved 25%–50% as submitted; and 6% approved more than 50% as submitted. *Id.* at 61. Lack of data about the IRB system is a notorious and persistent problem. *See, e.g.*, PRESIDENTIAL COMM'N FOR THE STUDY OF BIOETHICAL ISSUES, MORAL SCIENCE: PROTECTING PARTICIPANTS IN HUMAN SUBJECTS RESEARCH 33 (Dec. 2011); INST. OF MED., RESPONSIBLE RESEARCH: A SYSTEMS APPROACH TO PROTECTING RESEARCH PARTICIPANTS (Daniel D. Federman et al. eds., 2003); OFFICE OF THE INSPECTOR GEN., U.S. DEP'T OF HEALTH & HUMAN SERVS., OEI-01-97-00193 INSTITUTIONAL REVIEW BOARDS: A TIME FOR REFORM (1998) [hereinafter HHS REPORT]; Christine Grady, *Do IRBs Protect Human Research Participants?*, 304 J. AM. MED. ASS'N 1122 (2010); Thomson et al., *supra* note 35, at 96 n.6.

206. Hamburger, *supra* note 14, at 407, 424.

207. The Commission's survey of IRB practices during 1974–1975 found that, overall, IRBs required "modifications" in more than half of the protocols they reviewed. NAT'L COMM'N FOR THE PROT. OF HUMAN SUBJECTS OF BIOMEDICAL AND BEHAVIORAL RESEARCH, INSTITUTIONAL REVIEW BOARDS: REPORT AND RECOMMENDATIONS 60 (1978). However, the vast majority of these involved IRB requests for additional information from researchers or required changes to informed consent forms. Substantive changes were required in only 12% to 16% of reviewed proposals—3% to 4%, each, for changes to scientific design, subject selection, risks and discomforts, and confidentiality. *Id.* at 60–61. Like the *Bell Report*, the Commission found marked variability among IRBs, with 14% requiring "modifications" to all proposals and 22% requiring changes to no more than one-third of them. *Id.* at 61. *See also* Sue Richardson & Miriam McMullan, *Research Ethics in the UK: What Can Society Learn from Health?*, 41 SOCIOLOGY 1115, 1122 (2007) (51% of surveyed social scientists in the U.K. reported having been required by IRBs to alter their research for the worse at least once within the prior five years).

208. *See supra* note 64 and accompanying text.

1. To Researchers and Society

Researchers, of course, suffer when research is blocked, altered, delayed or abandoned.[209] Their careers can be diminished, stalled, or even ruined. Those who lack power (such as untenured faculty and students) or are on deadline (for graduation, tenure, or a grant) are especially vulnerable. Undergraduates and even master's level graduate students increasingly graduate without any experience in empirical research,[210] to the detriment of both them and society.

Society, too, clearly suffers from IRB risk aversion. Even acknowledging that much research ultimately fails to yield important knowledge, as Philip Hamburger notes,

> [O]ne only has to assume that one in 100,000 projects would otherwise have had profound benefits to understand the loss caused by IRBs. A single educational, epidemiological, or medical study can transform the lives of countless individuals—whether by giving rise to a lifesaving invention or treatment or by prompting a new government policy . . . .[211]

Nor do these numbers reflect the considerable research that is abandoned or foregone due to the chilling effect of IRB review. Ironically, the most innovative research tends to be the most likely to be blocked or foregone, since it often relies on unorthodox methodologies or involves cutting-edge topics likely to be viewed as controversial or "sensitive" by IRBs.[212]

2. To Prospective Participants

Less obviously—but critically—IRB risk aversion sets back the interests of those who might have chosen to accept an offer to enroll in a study but who, in the wake of an unfavorable risk–benefit decision by an IRB, receive a substantively different (and perhaps less attractive) offer than they otherwise would have had—or no offer at all. It is important that

209. *See, e.g.*, Jack Katz, *Toward a Natural History of Ethical Censorship*, 41 L. & SOC'Y REV. 797, 804 (2007) (researchers abandon plans for field research in favor of studying existing data). There is even some evidence that this chilling effect has spilled over to third parties. Grant reviewers, for instance, have confessed to assigning lower scores than they otherwise would to proposals that they predict may not survive IRB review.

210. *See* Richardson & McMullan, *supra* note 207 (discussing social science master's students in health in the U.K.).

211. Hamburger, *supra* note 14, at 469.

212. *See, e.g.*, Malcolm M. Feeley, *Legality, Social Research, and the Challenge of Institutional Review Boards*, 41 L. & SOC'Y REV. 757, 769–70 (2007); Annette Hemmings, *Great Ethical Divides: Bridging the Gap Between Institutional Review Boards and Researchers*, 35 EDUC. RESEARCHER 12, 15 (2006).

prospective participants understand that the primary goal of the participant–researcher relationship, unlike the patient–physician or client–practitioner relationship, is not to pursue the individual participant's best interests. But it does not follow from the fact that research is not *intended by the researcher* to serve individual participant interests that no individual participant can ever (rationally) view participation as *in fact* in her best interests, broadly construed, relative to her other alternatives. The vast majority of agreements between individuals are not intended by both parties to serve the best interests of one of them; rather, both parties typically expect to benefit from the arrangement, often in different ways and to different degrees, commensurate with their individual preferences and alternatives. The conventional wisdom of research regulators and ethicists, which holds that protecting participants primarily means protecting them *from* research, denies this reality.[213]

Recall the study in which researchers asked female undergraduates about various sensitive topics, and found that those who had experienced child abuse were both more likely to report distress due to remembering the past *and* more likely to report that participation had benefitted them than were other respondents.[214] Ninety-five percent of participants reported a positive cost/benefit ratio and agreed with the statement, "Had I known in advance what participating would be like for me I still would have agreed."[215] Another study found that 92% of participants reported that they would participate in the study again.[216] Yet, under their broad mandate to consider all conceivable risks, IRBs will almost certainly count the risk that participants will be distressed, and apply it to all prospective participants, while under their unduly narrow sense of what counts as a benefit, they will almost certainly not count these potential psychological benefits for any participants.

Or consider an example from biomedical research. In one study, 65% of terminally ill cancer patients participating in a Phase I drug trial predicted that they would receive psychological benefits[217] from the reassurance of routine physician contact during a time of uncertainty, the ability to exercise willpower in a situation otherwise marked by factors outside of their control, and the knowledge that they are helping future cancer patients. Yet many argue that Phase I oncology trials have an

213. I develop this argument in *Research Contracts*. *See generally* Meyer, *supra* note 22, at 4.

214. *See* text accompanying note 151.

215. Josef I. Ruzek & Douglas F. Zatzick, *Ethical Considerations in Research Participation Among Acutely Injured Trauma Survivors: An Empirical Investigation*, 22 GEN. HOSP. PSYCHIATRY 27, 30 (2000).

216. Hebenstreit & DePrince, *supra* note 149, at 66–67.

217. Agrawal & Emanuel, *supra* note 18, at 1077.

unacceptable risk–benefit ratio, in part because critics undercount expected therapeutic benefits and demand more favorable risk–benefit profiles from oncology research than from virtually identical chemotherapy agents used outside of trials,[218] but also because they ignore these psychosocial benefits of trial participation. Ironically, although we have bioethics to thank for emphasizing the intensely personal nature of how individuals die, in subjecting one option for dealing with terminal illness to a risk–benefit analysis that refuses to consider any inputs other than those that sound in extended lifespan, we also have bioethics to thank for restricting choice in dying.[219]

Given that IRBs can only make one risk–benefit decision binding on everyone, setting the risk–benefit profile to reflect the interests of the most vulnerable prospective participant might be said to be a prudent strategy. Indeed, it might even be said to be a requirement of justice. But, from the perspective of prospective participants—which is, after all, what the IRB system is designed to protect—this strategy is benign only if participants' welfare can only be set back, and never advanced, by research participation. The evidence suggests that this is not the case.

IV. PROPOSALS TO REFORM IRBS AND WHY THEY WILL NOT SOLVE (AND MAY EXACERBATE) THE HETEROGENEITY PROBLEM

Rather than responding to participant heterogeneity with risk aversion, can IRBs be reformed to permit them to accommodate prospective participant heterogeneity in their risk–benefit decisionmaking? There is no shortage of proposals to reform the IRB system, including those that would "centralize, regionalize, or consolidate IRBs, strengthen and demystify federal oversight, infuse more support and resources into the system, augment IRB member training, require credentialing of IRB professionals, mandate independent accreditation, educate the public, and continue to investigate 'alternative' models of review."[220] Most of these proposed reforms ultimately seek to redress what their proponents see as IRB risk–benefit "errors" (whether of the Type I or the Type II variety), for which IRB variation in risk–benefit analysis is the most salient symptom. Numerous empirical studies that have consistently found that IRBs—both

218. *See id.*

219. *Accord id.* ("It would be ironic if critics of phase 1 cancer studies considered only the physical benefits and ignored these quality-of-life and psychological benefits because they want to ensure a quality dying process for terminally ill patients.").

220. Lura Abbott & Christine Grady, *A Systematic Review of the Empirical Literature Evaluating IRBs: What We Know and What We Still Need to Learn*, J. EMPIRICAL RES. ON HUM. RES. ETHICS, Mar. 2011, at 3.

in the U.S.[221] and abroad,[222] and across a range of institutional settings and research methodologies—reach markedly different decisions about the acceptability of similar and even identical studies. Proposals differ chiefly in what they identify as the root cause of IRB risk–benefit error. Those that identify this cause as insufficient lay input into IRB risk–benefit decisionmaking and seek to increase such input in various ways. Other proposals view the ad hoc nature of IRB risk–benefit analysis as the primary problem and so seek redress by formalizing, or adding rigor to, IRB decisionmaking.

This Part considers these two sets of proposals and concludes that, although some reforms might yield independent benefits, few would significantly address the heterogeneity problem, none would solve it, and some might well exacerbate it.

---

221. *See supra* text accompanying notes 76–78, 115–118, 204; Ceci et al., *supra* note 72, at 1000; Kathleen Dziak et al., *Variations Among Institutional Review Board Reviews in a Multisite Health Services Research Study*, 40 HEALTH SERVS. RES. 279, 280 (2005); Lee A. Green et al., *Impact of Institutional Review Board Practice Variation on Observational Health Services Research*, 41 HEALTH SERVS. RES. 214, 214 (2006); Sarah M. Greene et al., *Impact of IRB Requirements on a Multicenter Survey of Prophylactic Mastectomy Outcomes*, 16 ANNALS EPIDEMIOLOGY 275, 276–78 (2006); Jerry Goldman & Martin D. Katz, *Inconsistency and Institutional Review Boards*, 248 J. AM. MED. ASS'N 197, 202 (1982); Jon Mark Hirshon et al., *Variability in Institutional Review Board Assessment of Minimal-Risk Research*, 9 ACAD. EMERGENCY MED. 1417, 1419 (2002); Jonathan Mansbach et al., *Variation in Institutional Review Board Responses to a Standard, Observational, Pediatric Research Protocol*, 14 ACAD. EMERGENCY MED. 377, 377 (2007); Rita McWilliams et al., *Problematic Variation in Local Institutional Review of a Multicenter Genetic Epidemiology Study*, 290 J. AM. MED. ASS'N 360, 365 (2003); Craig D. Newgard et al., *Institutional Variability in a Minimal Risk, Population-Based Study: Recognizing Policy Barriers to Health Services Research*, 40 HEALTH SERVS. RES. 1247, 1248 (2005); Henry Silverman et al., *Variability Among Institutional Review Boards' Decisions Within the Context of a Multicenter Trial*, 29 CRITICAL CARE MED. 235, 235 (2001); Thomas O. Stair et al., *Variation in Institutional Review Board Responses to a Standard Protocol for a Multicenter Clinical Trial*, 8 ACAD. EMERGENCY MED. 636, 637 (2001); Mary Terrell White & Jennifer Gamm, *Informed Consent for Research on Stored Blood and Tissue Samples: A Survey of Institutional Review Board Practices*, 9 ACCOUNTABILITY RES. 1, 1–2 (2002).

222. *See supra* text accompanying 119–127, 177–80; Hans-Peter Graf & Dennis Cole, *Ethics-Committee Authorization in Germany*, 21 J. INST. MED. ETHICS 229 (1995); Claire Gilbert et al., *Diversity in the Practice of District Ethics Committees*, 299 BMJ 1437, 1438 (1989); Ursula J. Harries et al., *Local Research Ethics Committees: Widely Differing Responses to a National Survey Protocol*, 28 J. ROYAL C. PHYSICIANS LONDON 150, 150 (1994); Matthew Hotopf et al., *Are Ethical Committees Reliable?*, 88 J. ROYAL SOC'Y MED. 31, 32 (1995) (U.K.); Jennifer Marshall & Michael R. Hadskis, *Canadian Research Ethics Boards, MRI Research Risks, and MRI Risk Classification*, IRB: ETHICS & HUM. RES., July–Aug. 2009, at 9, 14; Claire Middle et al., *Ethics Approval for a National Postal Survey: Recent Experience*, 311 BMJ 659, 660 (1995) (U.K.); G. Moutel et al., *Analysis of a Survey of 36 French Research Committees on Intracytoplasmic Sperm Injection*, 351 LANCET 1121, 1123 (1998).

A. Increased Lay Input into IRB Decisionmaking

The premise of the first set of reform proposals is that IRB risk–benefit error is primarily caused by insufficient lay input into IRB decisionmaking. Both camps of IRB critics have remarked that IRB decisions often poorly reflect participant interests despite the fact that IRBs' primary charge is to protect participant welfare. Recall, for instance, the "Type II error camp," for whom IRBs most often, or most seriously, err in permitting unreasonably risky research to proceed. This camp often attributes such errors to an IRB composed of conflicted or captured members whose incentives lie in protecting the institution or their fellow researchers (in the case of academic IRBs) or in appeasing the entity that pays them (in the case of commercial IRBs), rather than in protecting unidentified participants.

A "Type I error camp" version of this complaint arises cyclically as well, although it is generally reserved for clinical research investigating serious or life-threatening diseases. In the 1980s, for instance, various activists and patient advocates, particularly in the HIV/AIDS and breast cancer "communities," argued that they were being unreasonably denied opportunities to participate in clinical research by those who were effectively "protecting them to death." They sought access to research, and sometimes also demanded a role in the development of research agendas, study designs, and drug approval processes.[223]

1. Lay Membership on IRBs

One commonly proposed solution to such problems is to amend the regulations that govern IRB membership so that IRBs better reflect participants' interests. The architects of the current system deliberately chose to make IRBs local, rather than regional or national, so that they would reflect the community's values and other "local knowledge."[224] The National Commission, moreover, called for "diverse membership" on IRBs to reflect "awareness and appreciation of the various qualities, values and needs of the diverse elements of the community served by the institution or in which it is located."[225] The codification of this sentiment in the regulations is the requirement that each IRB "be sufficiently qualified

223. *See, e.g.*, Herbert R. Spiers, *Community Consultation and AIDS Clinical Trials: Part I*, IRB: ETHICS & HUM. RES., May–June 1991, at 7, 8. Today, disease communities are as likely to sponsor or conduct research themselves as they are to take up picket signs.

224. *See* NAT'L COMM'N, IRBS, *supra* note 24, at 1–2. Thus, although legal scholars often point to inconsistent IRB results, the regulatory framework embraces local variation as a strength of the IRB system rather than a problem to be solved.

225. *Id.* at 14.

through the experience and expertise of its members, and the diversity of the members, including considerations of race, gender, and cultural backgrounds and sensitivity to such issues as community attitudes, to promote respect for its advice and counsel in safeguarding the rights and welfare of human subjects."[226]

A large gap separates this aspiration from reality. First, the decision to establish IRBs at the local level predates globalization and the emergence of widespread, inexpensive telecommunications. Research today is very often conducted at multiple sites that span states and even countries, and much additional research is conducted online without respect to geographic boundaries. Under these circumstances, there is no particular reason to think that institution-based IRBs are especially representative of the population of prospective participants.

But IRBs are unlikely to reflect prospective participant preferences even when research is to be conducted at the same institution whose IRB reviews it. IRBs are by regulation required to have at least five,[227] and in practice have on average about thirteen,[228] members. IRBs are only required to include one member "whose primary concerns are in nonscientific areas" and one "who is not otherwise affiliated with the institution."[229] Often, the IRB finds one person to fulfill both roles. On average, then, one—or at most two—voting members on a thirteen-member IRB tend to be something other than scientific research faculty affiliated with the institution.

Moreover, neither the nonscience member nor the non-affiliate member need be a research participant, an advocate for participants, or even a random member of the local geographic community. Rather, the nonscience member might be (and often is) a humanities or law faculty member, while the non-affiliate might be a member of the research faculty at the institution across the street or, at best, a professional member of the community, such as a practicing attorney or clinician.

It should therefore not be surprising that, as a whole, IRB members are white, well-educated, institutionally affiliated, medically trained, and male.[230] In most cases, even the "non-scientist" and "non-institutional

226. 45 C.F.R. § 46.107(a) (2012). IRBs must also make "[e]very nondiscriminatory effort" to ensure that members include both men and women. *Id.* § 46.107(b).

227. *Id.* § 46.107(a).

228. Raymond G. De Vries & Carl P. Forsberg, *What Do IRBs Look Like? What Kind of Support Do They Receive?*, 9 ACCOUNTABILITY RES. 199, 203 (2002).

229. 45 C.F.R. § 46.107(c)–(d). The Commission, by contrast, had recommended that "at least one-third but no more than two-thirds of the IRB members should be scientists." NAT'L COMM'N, IRBS, *supra* note 24, at 14.

230. *See* BELL ET AL., *supra* note 45, at 23–24 (finding that 95% of surveyed IRB chairs

affiliate" members more closely resemble other IRB members than they do the "lay community" or the population of prospective research participants.[231] Nor should it be particularly surprising that token nonscience and non-affiliate members view their contributions to IRB deliberation as modest at best.

In response, many have argued for an increase in lay membership on IRBs,[232] or for a separate process of community consultation in some or all research.[233] But participant heterogeneity is far too fine-grained for these proposals to provide IRBs with significantly better information about prospective participant preferences than they currently have. Although geographic, special interest, and other communities are often constitutive of individual identity, each individual is simultaneously a member of multiple different communities. Within any of these communities, variation in research risk–benefit preferences will be significant at the unit of the individual, a fact suggested by the oft-noted difficulties in determining who is entitled to speak for a community, and even in identifying the relevant "community" itself. In short, proposals to enhance the local nature of IRB review by adding lay members are not local *enough* to solve the information aspect of the heterogeneity problem. And they would do nothing to address IRBs' aggregation problem.

2. Public Transparency and Accountability

IRB decisionmaking, like that of juries, is essentially conducted in the

---

were white, 77% were male, and 62% were full-time faculty affiliates of the institution); De Vries & Forsberg, *supra* note 228, at 202, 205–06 (survey of random sample of 10% of IRBs registered with OHRP (n=87) finding that: 85% of IRB members (n=1161) are white, with 28% of IRBs having all white members; in over 85% of IRBs, the majority of members are institutional affiliates).

231. *See, e.g.*, Emily E. Anderson, *A Qualitative Study of Non-Affiliated, Non-Scientist Institutional Review Board Members,* 13 ACCOUNTABILITY RES. 135, 140, 151 (2006) (small study (n=16) of non-affiliated and nonscientist IRB members finding most were white, educated, professional, had no experience as research participants, and were recruited through institutional contacts).

232. *See, e.g.*, NBAC, *supra* note 26, at xvi (recommending that at least 25% of each IRB's members be nonscientists, non-institutional affiliates, or participant advocates); HHS REPORT, *supra* note 205. at 17–18 (recommending increased nonscientist and noninstitutional membership as a means of mitigating IRB capture by institutional interests); Jesse A. Goldner, *An Overview of Legal Controls on Human Experimentation and the Regulatory Implications of Taking Professor Katz Seriously*, 38 ST. LOUIS U. L.J. 63, 107 (1993) (laypersons constitute at least half of IRB members in Denmark).

233. *See, e.g.*, Neal Dickert & Jeremy Sugarman, *Ethical Goals of Community Consultation in Research*, 95 AM. J. PUB. HEALTH 1123 (2005) (discussing the ethical goals that community consultation achieves).

proverbial black box. It is barely transparent to researchers, much less to the public. IRB meetings are generally closed to researchers. And although agencies have access to the minutes of IRB meetings, which must include "the basis for requiring changes in or disapproving research" and a "written summary of the discussion of controverted issues and their resolution,"[234] these documents generally are not deemed public records under the Freedom of Information Act (FOIA). The regulations require IRBs to provide researchers with a written "statement of the reasons for its decision" whenever it disapproves research, and to allow researchers "an opportunity to respond in person or in writing."[235] But this statement need not be, and in practice rarely is, detailed, and no explanation at all is required for IRB decisions requiring substantive alterations to research—almost certainly the much larger category.

IRBs could, of course, be made to publicly account for their decisions, as a matter of either law or private policy. IRB minutes could be posted online on institutional or agency websites (with redaction as necessary to protect confidential information); as with FOIA requests, made available to anyone upon request;[236] or included with any published results of the research they have reviewed.[237] But these reforms would only permit comment on IRB decisions after the fact, when any public commentary would come too late, at least for the study at bar. And some of these reforms, such as tying IRB transparency to research publication, fail to address IRB decisions regarding research that is not published or is rejected by IRBs.

To allow for more timely public input, IRBs might be required to hold meetings open to the public (again, with closed door proceedings permitted when necessary to ensure confidentiality). The regulations do not provide researchers, much less potential participants, with a right of appeal, and institutions rarely voluntarily provide such processes. But that, too, could be changed. The real difficulty, instead, is one that is also reflected in the relative failure of public comment periods during administrative rulemaking: even assuming that their voice might have an impact, very few members of the public have sufficient incentive to monitor or comment on IRB decisions. The benefit that an individual can expect from participating in research will usually be modest relative to the effort she would have to

234. 45 C.F.R. § 46.115(a)(2), (b).

235. *Id.* § 46.109(d).

236. *See, e.g.*, MD. CODE ANN., HEALTH-GEN. § 13-2003 (West 2012) (requiring IRBs to make minutes available to any person upon request).

237. *See* Alexander Halavais, *Social Science: Open Up Online Research*, 480 NATURE 174, 175 (2011) (arguing that funders and journals condition their services on IRB agreement to provide their "ethics reflections" for the benefit of researchers and other IRBs).

exert to monitor IRB decisionmaking. Nor, again, would public transparency address the aggregation facet of the heterogeneity problem.

In short, although greater IRB transparency and accountability is independently desirable, it, like increased lay membership on IRBs, cannot solve the heterogeneity problem. Moreover, if taken too far, this line of reform could exacerbate the heterogeneity problem. Many, for instance, advocate a court-like hierarchy of local, regional, and national IRBs in which decisions by "lower" IRBs are subject to appellate review by "higher" IRBs.[238] Like judicial systems, this mechanism would result in not only intra- but also inter-IRB consistency through precedent. But where participant heterogeneity obtains, an increase in top-down, control-and-command risk–benefit analysis might make things worse.

B. Rigorous, Evidence-Based IRB Decisionmaking

A second set of proposals is based on the premise that the root cause of IRB risk–benefit error is insufficiently rigorous or evidence-based IRB decisionmaking. In its discussion of IRB risk–benefit analysis, the Commission acknowledged that "[t]he possible harms and benefits from proposed research involving human subjects may not be quantifiable," but nevertheless insisted that they "should be evaluated systematically to assure a reasonable relation between" them.[239] In practice, however, IRB risk assessment "is highly unstructured; essentially, the members are simply given a set of protocols and asked for their reactions."[240] Indeed, a study of fifty-three IRBs in the Netherlands found that only six used anything like a systematic approach to risk–benefit analysis. The others admitted to relying solely on intuition.[241]

1. Improved IRB Risk–Benefit Methodology

Some thus argue that the solution to IRB risk–benefit "error" is a more systematic methodology for assessing research risks (research benefits are essentially ignored). For instance, some advocate for increased staffing,

238. Coleman, *supra* note 12, at 43.

239. NAT'L COMM'N, IRBS, *supra* note 25, at 23. The Commission also noted that risk–benefit "evaluation should include an arrayal of alternatives to the procedures under review and the possible harms and benefits associated with each alternative." *Id.* IRBs generally do not even pretend to speculate about the risks and expected benefits of alternatives to research participation, nor do proposed methodologies for IRB risk–benefit analysis involve such comparisons.

240. *See* Coleman, *supra* note 12, at 14; Rid et al., *supra* note 97, at 1472.

241. H.E.M. van Luijn et al., *Assessment of the Risk/Benefit Ratio of Phase II Cancer Clinical Trials by Institutional Review Board (IRB) Members*, 13 ANNALS ONCOLOGY 1307 (2002).

training, or funding of IRBs. But disparities among IRBs in staffing, training, and funding do not seem to correlate with, much less cause, IRB variation.

Others urge IRBs to issue court-like opinions, pointing to the potential debiasing effect of such opinions on otherwise ad hoc risk–benefit reasoning: "Because [IRB] members are not required to state reasons for their decisions, the process encourages reliance on impressionistic judgments, or 'gut reactions.'"[242] Forcing IRBs to specify reasons for their decisions, they say, may curb their tendency to rely on indefensible gut reactions and should help ensure consistency across protocols and time within IRBs. But if participant heterogeneity is real, then formal IRB risk–benefit decisions may be based on post-hoc rationales to justify inescapably subjective decisions. And precedent-based IRB decisionmaking—like other attempts at increasing consistency among IRBs[243]—would only magnify the heterogeneity problem.

Still others advocate thinking systematically about research risk in comparison to other risky activities.[244] In this vein, research participation has been (or could be) compared to risky employment (e.g., firefighting, law enforcement, emergency rescue work,[245] logging, mining, commercial fishing, armed services, professional football, test piloting), risky leisure

242. Coleman, *supra* note 12, at 14.

243. For example, some argue that IRB variation is caused by excessive allowance for IRB discretion, which could be limited through clearer, more specific risk–benefit regulations or agency guidance. *See, e.g.*, de Champlain & Patenaude, *supra* note 177, at 533–34; Lenk et al., *supra* note 18, at 87 (arguing that variation in IRB risk analysis should be addressed through "more detailed, comprehensive, and unambiguous regulation . . . that does not permit such a wide range of interpretation").

244. Some of the riskiest activities are relatively unregulated, compared to research. These include sports, leisure, and other voluntary activities such as one pack-a-day smoking (annual risk of death: 300 per 100,000 persons at risk), parachuting (200), motorcycling (65), skydiving (58), hang gliding (26) and boating (5), as well as occupations such as lumberjack (118), farmer (28), miner (27), police officer (20) and firefighter (10). Aaron Wildavsky & Adam Wildavsky, *Risk and Safety*, *in* THE CONCISE ENCYCLOPEDIA OF ECONOMICS tbl.2 (2008), *available at* http://www.econlib.org/library/Enc/RiskandSafety.html. On the other hand, much research more closely resembles risky activities that are totally unregulated, not even by mandatory disclosure rules. This includes, for example, survey research and polling; interview-based research and journalism or ordinary conversing.

245. Alex John London, *Two Dogmas of Research Ethics and the Integrative Approach to Human-Subjects Research*, 32 J. MED. PHIL. 99, 115 (2007); *see also* Benjamin I. Sachs, *Enabling Employee Choice: A Structural Approach to the Rules of Union Organizing*, 123 HARV. L. REV. 655, 681 (2010) (discussing the potential risks faced by employees who support a unionization effort); Alex John London, *Reasonable Risks in Clinical Research: A Critique and A Proposal for the Integrative Approach*, 25 STAT. MED. 2869, 2881 (2006) (research risks "must not be greater than the risks . . . permitted in . . . other socially sanctioned activities that are similar in structure to the research enterprise").

activities (e.g., skydiving, mountain climbing, bungee jumping, extreme martial arts), and risky charitable activities (e.g., living organ or tissue donation,[246] disaster relief work, volunteer work of various kinds in hostile territories, donating money to charity despite a volatile economy and employment sector). These comparisons might be used either to classify strata of research risks or to define acceptable risk limits.

What such compare-and-contrast exercises mostly suggest, however, is the exceptional way we manage research risks.[247] Normatively speaking, little if anything follows from these observations.[248] More to the point, preferences regarding these comparator activities are also heterogeneous. When the Environmental Protection Agency surveyed "a range of health risks" along with government and private standards of acceptable risk in order to glean therein some level of "acceptable risk" in those areas which the agency might then use to inform its determination of acceptable levels of a particular pollutant, it found that "[n]o fixed level of risk could be identified as acceptable in all cases and under all regulatory programs." Rather, "the acceptability of risk is a relative concept and involves consideration of different factors," including the

> certainty and severity of the risk; the reversibility of the health effect; the knowledge or familiarity of the risk; whether the risk is voluntarily accepted or involuntarily imposed; whether individuals are compensated for their exposure to the risk; the advantages of the activity; and the risks and advantages for any alternatives.[249]

2. Evidence-Based Risk–Benefit Analysis and Risk-Proportionate Regulation

Finally, many argue that IRBs should base risk–benefit analysis on

246. *See* F.G. Miller & S. Joffe, *Limits to Research Risks*, 35 J. INST. MED. ETHICS 445, 447 (2009).

247. We permit competent individuals to assume greater risks in activities that often have far less social value than HSR (and may even involve negative externalities). And we do so on the basis of consent alone, usually accompanied by far weaker (if any) information disclosure, risk minimization, and liability waiver rules; in few, if any, of the above examples do we require anything approaching the kind of case-by-case licensing procedure to which researchers are subjected every time they wish to engage in that activity.

248. The bare fact that we treat research and other risks inconsistently, without (much) more, cannot tell us whether, on one hand, IRBs should be more lenient in vetting research risks or, on the other hand, the Occupational Safety and Health Administration should be less lenient in permitting employment risks, the missions of Medécins Sans Frontiers should be subject to independent prospective review of their risks and expected benefits, and third parties should debate the "reasonableness" of tithing or writing checks to particular charities before anyone is permitted to do so. *Accord* Miller & Joffe, *supra* note 246, at 446–47.

249. *See* Baruch Fischhoff, *Acceptable Risk: A Conceptual Proposal*, 5 RISK: HEALTH, SAFETY & ENV'T 1, 2–3 (1994).

evidence of research outcomes rather than on their "'gut' feeling."[250] Data regarding participant outcomes is not routinely collected, nor does there currently exist any formal mechanism through which IRBs might share such data with one another. Data about participants' actual experiences, these commentators say, would enable IRBs to manage the "difficult balancing act" of permitting important research to go forward, but "without harm or jeopardy to individual participants."[251] Such data might be routinely collected on research outcomes[252] and fed into a national database of research risks (and, one hopes, benefits), for incorporation into future IRB risk–benefit analysis.[253]

The same data might also inform risk-based stratification of research types at the regulatory or statutory levels. Many have distinguished physical and nonphysical risks,[254] with some proposing to make research

250. Fendrich et al., *supra* note 73, at 35.

251. Decker et al., *supra* note 75, at 55.

252. The Commission itself had recommended that IRBs adopt procedures for continuing review of research, such as observing, "on a sample or routine basis," participant recruitment, the consent process, or the conduct of the research itself, as well as soliciting information from participants through interviews or feedback forms. NAT'L COMM'N, IRBS, *supra* note 25, at 16–17; *see also* CTR. FOR ADVANCED STUDY, IMPROVING THE SYSTEM FOR PROTECTING HUMAN SUBJECTS: COUNTERACTING IRB "MISSION CREEP" 18 (2007) (recommending collection of data regarding, *inter alia*, "what subjects perceive as risk, and what kinds of benefits to subjects and their communities make the relationship fair"); NAT'L RESEARCH COUNCIL, *supra* note 43, at 159–63 (recommending that researchers publish data "about types, incidence, and magnitude of harm encountered in social, behavioral, and economic sciences research" derived through debriefing their participants, and that OHRP "establish an ongoing system for collecting and publishing data that can help assess how effectively IRBs protect human research participants, how efficiently they review research, and how commensurate review is with risk"); David Wendler et al., *Quantifying the Federal Minimal Risk Standard: Implications for Pediatric Research Without the Prospect of Direct Benefit*, 294 J. AM. MED. ASS'N 826 (2005) (discussing the need for empirical data to guide IRB review); Brian Mustanski, *Ethical and Regulatory Issues with Conducting Sexuality Research with LGBT Adolescents: A Call to Action for a Scientifically Informed Approach*, 40 ARCHIVES SEXUAL BEHAV. 673, 674 (2011) (highlighting the need to change the IRB process of risk–benefit analysis from being subjective to being evidence-based).

253. *See, e.g.*, Annette Rid & David Wendler, *A Proposal and Prototype for a Research Risk Repository To Improve the Protection of Research Participants*, 8 CLINICAL TRIALS 705 (2011). Such proposals rarely think to include data about the benefits of research participation, but we can easily imagine including such data as well.

254. *See, e.g.*, Oakes, *supra* note 48, at 449 ("We know a fair amount about . . . physical risks . . . ; the IRB system was set up to address these. We know little about . . . nonphysical risks . . . , and this creates problems. How do we measure and weigh an annoying journalistic inquiry? What about a threat to confidentiality? Deception? Sensitive questions about illegal drug use?"); Mustanski, *supra* note 252, at 680 ("In almost all cases of social/behavioral research, [IRB risk–benefit analysis] involves a subjective determination based on opinions about the probability of a risk outcome occurring and its likely

review less burdensome by limiting IRB review to "high-risk" HSR involving physical interventions while deregulating "low-risk" HSR involving "mere" psychosocial risks.[255]

In July 2011, HHS and the Office of Science and Technology Policy released *Human Subjects Research Protections: Enhancing Protections for Research Subjects and Reducing Burden, Delay, and Ambiguity for Investigators*, an Advance Notice of Proposed Rulemaking (ANPRM) that contemplates potentially

---

consequences.").

255. *See, e.g.*, ZACHARY M. SCHRAG, ETHICAL IMPERIALISM: INSTITUTIONAL REVIEW BOARDS AND THE SOCIAL SCIENCES 1965–2009 (2010) (suggesting that biomedical and behavioral (i.e., psychology) research be regulated differently than social science, humanities and education research, primarily on the basis of differences in riskiness); Philip Hamburger, *The New Censorship: Institutional Review Boards*, 2004 SUP. CT. REV. 271, 272 (invoking, implicitly, the distinction by lamenting that a researcher must "get . . . prior permission not only if he wants to conduct a dangerous physiological experiment but also if he merely wants to ask individuals about their political opinions"); CTR. FOR ADVANCED STUDY, *supra* note 252, at 5 ("Those cases that pose the greatest chances for risk and harm are, and always have been, in the fields of biomedical research in general and clinical trials in particular."); E.L. Pattullo, *Commentary: Exemption from Review, Not Informed Consent*, IRB: ETHICS & HUM. RES., Sept.–Oct. 1987, at 6, 6 (proposing that prospective review be limited to research involving deception, intrusion upon the subject's person, or denial of or reduction in benefits); Oakes, *supra* note 48, at 449 fig. 1 (depicting "Typical Risk Spectrum by Research Discipline," according to which the low-risk end of the spectrum begins with journalism, whose primary risk is annoyance, proceeds through oral history, anthropological investigations, evaluation research, biomedical epidemiological studies and pharmacological trials, and concludes at the high-risk end with surgical trials, which carry the risk of death).

Often, the proposed regulatory distinction between physical and nonphysical risks is, in turn, mapped onto a second distinction, between biomedical and nonbiomedical (i.e., social science, educational, humanities, and perhaps behavioral) research. The facts that the scandals that led to the current governance regime almost exclusively involved biomedical research and that the National Research Act refers repeatedly to "biomedical and behavioral research," with no clear mention of social science, and no mention whatever of education or humanities research, do make for odd beginnings in a story that ends with the broadly applicable regulations described in Part I. Typically, the proffered standard relies on a distinction between the psychosocial risks that are typical of nonbiomedical research and the physical risks that are typically limited to biomedical research.

But questions of agencies' statutory authority to apply the regulations widely aside, risk type does not neatly correspond to research methodology. All methodologies and all disciplines have the potential to set back the interests of research participants. Biomedical studies often involve no serious physical risks, or even any physical risks at all, but sometimes do involve psychosocial, economic, or legal risks. Conversely, behavioral and social science studies can involve physical risks.

We might abandon methodological proxies and base risk-proportionate regulation on the kinds of risks themselves—physical versus nonphysical. But physical harms are not, as a rule, always greater in magnitude, more costly for the victim or society, more irreversible, or otherwise more important to avoid than are psychological, social, economic, and legal harms.

sweeping changes to the regulations under which IRBs operate.[256] As its title suggests, the ANPRM attempts to appease both camps of critics by reallocating IRB review and agency oversight resources from "low-risk" studies, where they are inefficient, to studies that "pose risks of serious physical or psychological harm," which currently suffer from insufficiently rigorous review.

The ANPRM is but one nation's contribution to a global trend toward "risk-proportionate" regulation of HSR,[257] increasingly supported by scholars,[258] in which the kind and extent of IRB review and other oversight are proportionate to the riskiness of the research. Risk-proportionate regulation has the potential to appease both camps by IRB reallocating resources that are today spent on review of low-risk research to a more thorough review of high-risk research, thus simultaneously eliminating unnecessary burdens on benign research and freeing up resources to better protect participants from serious harms. It aims for two politically unassailable goals—the safety and welfare of research participants and the efficient use of scarce resources—and wraps these goals in the seemingly unobjectionable language of "proportionality."

Of course, risk-proportionate regulation of HSR requires a meaningful way for IRBs, regulators or legislators to distinguish "low-" from "high-risk" research. That is, it requires a basis on which some social planner can, in advance and with respect to all prospective participants, deem some research-related harms insufficiently probable or significant to warrant the full panoply of protections afforded participants in other studies. Using research outcomes data as the basis for risk-proportionate regulation (or for "evidence-based" IRB risk–benefit analysis at the protocol

256. ANPRM, *supra* note 34.

257. *See id.*; *see also* Eur. Comm'n on Health and Consumers Directorate-Gen., *Revision of the 'Clinical Trials Directive', 2001/20/EC*, SANCO/C/8/PB/SF D(2011) 143488 (Sept. 2, 2011); ACAD. OF MED. SCIS., A NEW PATHWAY FOR THE REGULATION AND GOVERNANCE OF HEALTH RESEARCH (2011) (U.K.); PRESIDENTIAL COMM'N FOR THE STUDY OF BIOETHICAL ISSUES, *supra* note 205, at 33 (U.S.); NAT'L RESEARCH COUNCIL, *supra* note 43, at 143 (noting that "IRBs in some instances may overestimate the risks of harm to participants in [social, behavioral, and economic sciences] (and biomedical) research" while "[a]t the other extreme a few IRBs may underestimate risk" and calling for IRB review that is commensurate to study risk); INST. OF MED., *supra* note 205.

258. *See, e.g.*, Dale Carpenter, *Institutional Review Boards, Regulatory Incentives, and Some Modest Proposals for Reform*, 101 NW. U. L. REV. 687, 688–89 (2007) (proposing that IRBs be required to expedite review of low-risk social science research, absent a finding that the study's risks "substantially outweigh" its anticipated benefits) (emphasis omitted)); Scott Kim, Peter Ubel & Raymond De Vries, *Pruning the Regulatory Tree*, 457 NATURE 534 (2009) (suggesting that low-risk research be exempt from IRB review); Adil E. Shamoo, *Deregulating Low-Risk Research*, CHRON. REV., Aug. 3, 2007, at B16 (supporting the exemption of low-risk studies from federal regulation).

level) faces significant problems. In addition to the not inconsiderable expense involved,[259] such data would address the information aspect of the heterogeneity problem only crudely. Data about the preferences of particular individuals based on debriefing from particular past studies can only very imperfectly predict the preferences of other, as-yet unidentified individuals regarding other, future studies. And, like other reform proposals, evidence-based risk–benefit analysis would do nothing to address the aggregation problem.[260]

CONCLUSION

IRBs are quasi-governmental actors charged with protecting research participants, largely by permitting participants to be invited to enroll in a study if and only if the IRB determines that the study's risks to participants are reasonable in relation to its expected benefits for them, if any, and for society. But as central actors who lack information about the preferences and circumstances of individual prospective participants—and who, in any case, must make a single decision for each study applicable to diverse individuals—IRBs are incapable of determining whether a study's risk–benefit profile is in fact "reasonable" to any particular prospective participant. Moreover, by employing a broad understanding of research risk and a narrow understanding of research benefit, IRBs tend to assign to proposed studies risk–benefit profiles that are likely to reflect the preferences of the most risk-averse minority of prospective participants. Erring on the side of restricting the kinds of research opportunities to which individuals may be invited is benign only if we assume that individuals can only be harmed, and never benefitted, by participating in research. That claim is refuted by much of the empirical evidence discussed in this Article.

Given its intractability, it should come as little surprise that, on the rare

259. The National Commission itself recognized that this massive data collection and analysis would entail "a substantial strain on the limited resources" of IRBs, which are now under exponentially greater constraints than they were in 1978. NAT'L COMM'N, IRBS, *supra* note 25, at 16–17. And since such data collection would itself constitute human subject research, still further IRB costs would presumably accrue.

260. Moreover, even if past preference data sufficiently reflected all prospective participants' preferences (in the highly unlikely case that participant preferences turn out to be both homogeneous and stable), there are reasons to doubt that IRBs would base their risk–benefit decisions on at least some of this data. IRBs, like all regulators (indeed, arguably like all human actors), have psychological incentives to resist data when it undermines their *raison d'être*—in this case, data that suggests that research entails very little risk. Similarly, considerable empirical research suggests that decisionmakers are more risk averse when deciding for others than they are when they make the same choice for themselves. The debiasing effect of data on this tendency is unclear. I discuss these and other regulatory and cognitive biases of IRBs in a companion article. *See* Meyer, *supra* note 22.

occasion when a regulator acknowledges the heterogeneity problem, its counsel to IRBs does not inspire confidence. New York State's HSR guidelines offer a representative example:

> [T]olerance for discomfort and inconvenience may vary considerably, causing what may be perceived as ordinary discomfort or inconvenience by some subjects to escalate to significant harm for others. Examples include a bronchoscopy or bone marrow biopsy, which may be experienced as unpleasant by some subjects and as severe discomfort by others, or a sleep study that reverses day and night, which upon completion may require no readjustment by some subjects and a psychologically difficult readjustment by others.[261]

New York IRBs are told to be "cognizant of" participant heterogeneity and to "consider this information in their assessment of the protocol's risks."[262] As we have seen, however, the idea that IRB risk–benefit analysis meaningfully establishes when research risks are reasonable in relation to their expected benefits is a legal fiction.

We can continue to perpetuate this legal fiction, demanding that quasi-governmental actors perform an essentially impossible task; alienating and disillusioning researchers; and driving researchers and IRBs to strategies of evasion and risk aversion, respectively, that are costly for all stakeholders, reducing the amount, quality, and timeliness of knowledge production, and—less obviously, but just as importantly—denying would-be participants valuable opportunities that would advance their welfare.

Or, we can embrace the implicit utilitarianism of those scholars of regulation who argue that the extent to which IRBs protect participants is not justified by the costs to researchers and society of prospective IRB review. That is, we could reconceive research as "reasonable" whenever its total expected benefits for society and participants (if any) *outweigh* its expected costs—that is, whenever research is expected to maximize social welfare. This, however, would constitute a radical departure from the historical purpose of our system of HSR governance. The Common Rule is singularly devoted to protecting participants from research-related harm. Just as neither Title II of the National Research Act nor its implementing regulations contemplates cost–benefit analysis or any other way in which individual participant welfare might be acceptably sacrificed in service of social welfare,[263] they do not contemplate knowingly sacrificing the welfare of participants with idiosyncratic preferences.

I suggest that there is a third way. We can embrace rather than deny

261. N.Y. IRB GUIDELINES, *supra* note 181, at 9 (footnote omitted).
262. *Id.*
263. *See supra* Part II.C.1.a.

participant heterogeneity, while at the same time increasing the quantity, quality, and speed of research, and remaining faithful to the values that have historically driven HSR governance: participant welfare and respect for participant autonomy. If IRBs cannot meaningfully factor participant heterogeneity into their risk–benefit assessment, rather than futilely tweaking this or other aspects of the IRB system, we should jettison centralized risk–benefit reasonableness inquiries and replace with a system of private ordering that is sensitive to heterogeneous preferences. Individual prospective participants, not IRBs, should decide whether it is reasonable *for them* to accept the risks of participating in a particular research study. I take up this alternative framework in a companion Article.[264]

264. *See* Meyer, *supra* note 22.