



HHS Public Access

Author manuscript

Account Res. Author manuscript; available in PMC 2017 September 28.

Published in final edited form as:

Account Res. 2017 ; 24(7): 407–432. doi:10.1080/08989621.2017.1363650.

Minimal Risk in Pediatric Research: A Philosophical Review and Reconsideration

John Rossi, V.M.D., M. Bioethics^a and Robert M. Nelson, M.D., Ph.D.^b

^aDornsife School of Public Health, Drexel University, Philadelphia, Pennsylvania, USA

^bOffice of Pediatric Therapeutics (OPT), Office of the Commissioner (OC), United States Food and Drug Administration (FDA), Silver Spring, Maryland, USA

Abstract

Despite more than thirty years of debate, disagreement persists among research ethicists about the most appropriate way to interpret the U.S. regulations on pediatric research, specifically the categories of “minimal risk” and a “minor increase over minimal risk.” Focusing primarily on the definition of “minimal risk,” we argue in this article that the continued debate about the pediatric risk categories is at least partly because their conceptual status is seldom considered directly. Once this is done, it becomes clear that the most popular strategy for interpreting “minimal risk”—defining it as a specific set of risks—is indefensible and, from a pragmatic perspective, unlikely to resolve disagreement. Primarily this is because judgments about minimal risk are both normative and heavily intuitive in nature and thus cannot easily be captured by reductions to a given set of risks. We suggest instead that a more defensible approach to evaluating risk should incorporate room for reflection and deliberation. This dispositional, deliberative framework can nonetheless accommodate a number of intellectual resources for reducing reliance on sheer intuition and improving the quality of risk evaluations.

Keywords

Ethics and public policy; history of research ethics; human subjects ethics; human subjects regulation and oversight; pediatric research ethics; research ethics

Introduction: The National Commission and the U.S. regulations on pediatric research

The current United States (U.S.) regulations on pediatric research were adopted a little over thirty years ago, after extensive deliberation by a national bioethics commission. In 1978, that commission—the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (hereafter “National Commission” or “Commission”)

CONTACT: John Rossi, jar444@drexel.edu, Dornsife School of Public Health, Drexel University, 3215 Market Street, MS457, Philadelphia, PA 19104.

Disclaimer

The views expressed in this article are solely those of the author (RMN) and do not represent the views of the Food and Drug Administration.

—published their report, *Research Involving Children* (1978). This report contains discussion about the fundamental ethical permissibility of pediatric research, particularly research not benefiting the child involved, as well as a series of specific recommendations that were first adopted—nearly verbatim—into the U.S. federal regulations in 1983 (US DHHS 1983, US DHHS 1991; Subpart D).

The U.S. pediatric research regulations draw an important ethical distinction between interventions (hereafter “research”)¹ that offer a prospect of direct benefit and interventions that offer no prospect of direct benefit. When a research intervention offers the prospect of direct benefit, it is permissible under the regulations if its risks are minimized as much as possible, if the anticipated benefit to the child justifies the risk, and if the risk-benefit profile is comparable to available alternatives (e.g., established treatments).

However, non-beneficial research (that is, research not offering a prospect of direct benefit to the child involved) conducted on healthy children is permissible under the regulations only when such research presents no more than minimal risk. “Minimal risk” is defined as risk where “the probability and magnitude of harm or discomfort anticipated in the research 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” (U.S. Department of Health and Human Services [US DHHS] 1991).² To clarify, the regulatory requirement here is not that the risks of research be identical to the risks of daily life activities or routine examinations; rather, daily life activities and routine examinations set a *level of equivalence* for research risks. The Commission’s reasoning in adopting this restriction was that children deserve special protection in research due to their vulnerability and inability to provide informed consent.

In addition to “minimal risk,” the National Commission and resultant regulations also include the category of “a minor increase over minimal risk.” This category was created as a standard for acceptable risk for non-beneficial research interventions or procedures in children with a “condition or disorder,” where the intervention or procedure is important to answering a research question concerning the condition or disorder (*Research Involving Children* 1978; DHHS 1991; Subpart D). “A minor increase over minimal risk” is not defined in the Commission’s report or in the regulations, and the National Commission’s ethical rationale for creating this category is not made entirely clear in their report. Finally, where non-beneficial pediatric research does not conform to either of these categories (e.g., presents more than a minor increase over minimal risk to children) but is nonetheless thought to be important and ethically defensible, it may be referred for review by a special federal panel (sometimes called “407 review” because this provision is located in 45 CFR 46.407) (U.S. Department of Health and Human Services [US DHHS] 1991 Subpart D).

¹While the term “research” appears in the title of the U.S. federal regulations pertaining to human-subjects research (45 CFR 46 and 21 CFR 50), and while the category of “minimal risk” refers to “research,” other categories of the pediatric regulations use the term “intervention.” The use of the term “intervention” highlights that the unit of analysis for ethical and regulatory purposes is not the research project as a whole, but rather each intervention that is to be employed in the course of the research.

²Though one might try to distinguish “low risk” from “minimal risk,” there is no indication in *Research Involving Children* that the National Commission thought of the two terms as distinct risk levels. Instead, they seemed to use “low” and “minimal” risk interchangeably and more-or-less synonymously—though “minimal risk” is the term that was chosen and defined in the regulations. For the purposes of this article, we treat the terms synonymously.

Henceforth, and for expediency, we refer to “a minor increase over minimal risk” and “more than a minor increase over minimal risk” as “derivative terms” of minimal risk.

Many if not most research ethicists seem to share the conviction that low-risk, non-beneficial pediatric research can be ethically defensible. Nonetheless, the decades since the introduction of the pediatric research regulations have witnessed a substantial amount of debate about how best to interpret minimal risk and its derivatives, with consensus still eluding the bioethical community (see, e.g., Binik 2014; Fisher, Kornetsky, and Prentice 2007; Freedman, Fuks, and Weijer 1993; Kopelman 2004a; Nelson 2011; Resnik 2005; Snyder, Miller, and Gray 2011; Wendler 2005; Westra et al. 2011). We believe that continued disagreement about the interpretation of minimal risk stems partly from the fact that fundamental philosophical issues concerning its definition and meaning are rarely considered directly. Many commentators advance their preferred interpretation of this term without first considering what kind of judgment minimal risk is supposed to represent or what kind of rational standards might apply to how we should make such judgments. As a result, it is not clear how interpretive disagreements should be resolved when they inevitably arise.

In this article, we probe the fundamental meaning of “minimal risk” and its derivative terms.³ We begin by reviewing debates to date and clarifying concerns at issue, since this will help to illuminate the problems to which we respond in subsequent sections of the article. We argue that minimal risk is best understood as a judgment of risk magnitude—being roughly synonymous with “low risk”—or, relatedly, as a judgment about “acceptably low risk,” and that on either interpretation the term is value-laden and not easily captured by rigid definitions. For this reason, the predominant argumentative strategy in the bioethical literature—defining minimal risk as a set of risks—is unlikely to result in epistemic closure of debates. We consider the kinds of rational constraints that might apply to evaluations of minimal risk, concluding that such evaluations are likely to be heavily intuitive in nature. However, resources are available to improve upon sheer intuition, and we conclude by sketching an alternative model for risk evaluation, the dispositional, deliberative model.

Debate over the U.S. pediatric research regulations: A brief review

Although it initially appears straightforward, the federal definition of “minimal risk” incorporates several ambiguities, the first of which arises because the definition uses both daily life and routine examination risks as a standard. There is no reason to expect the risks of daily life to be qualitatively similar or similar in magnitude to the risks of routine examinations, and the wording of the definition suggests that one can choose which standard to use. Second, reflection on the risks of daily life indicates that this category is ambiguous and requires specification. It could refer to the risks that every child faces in their daily life

³To clarify, our goal is not to argue whether the current regulatory structure is ethically defensible. Rather, our goal is to ask how the regulations should be interpreted and applied, consistent with how they are currently written, and what the risk categories contained in them are supposed to represent. For example, it might be argued that the category “a minor increase over minimal risk” should be eliminated because it is morally indefensible, since there is no good reason why we should expose children with a “condition or disorder” to greater risk than healthy children in research (see Ross 2003; Kopelman 2004b; for discussion). This question does not relate to the interpretation of “a minor increase over minimal risk” but instead to whether we should have the category in the first place.

(thus excluding risks not faced by all children); the risks any child faces in daily life (a much more permissive standard); the risks that the particular child being enrolled in research faces in their daily life (the so-called “relative” interpretation); the risks that “average, normal, healthy” (IOM 2004, p. 5) children face in their daily lives; or the daily life risks that are generally deemed socially allowable (see Kopelman 2004a).

Bioethicists typically attempt to resolve these ambiguities by stipulating a more concrete interpretation of minimal risk. Some have argued that “the risks of daily life” should only include those risks faced by “average, normal, healthy” (Institute of Medicine (IOM) 2004, p. 5) children. Others have argued that the “risks of daily life” clause should be eliminated entirely and that only the “routine examinations” interpretation should be used (Fisher, Kornetsky, and Prentice 2007; Resnik 2005). Still others have proposed new interpretations of minimal risk, such as the risks that children are allowed to encounter in charitable activities (Wendler 2005) or the daily life risks faced by children who are not unduly burdened (Binik 2014).

Consensus has unfortunately eluded these debates. Critics of the relative interpretation of “the risks of daily life” assert that it is unjust because it allows different children to be exposed to different levels of risk absent a good justification, the consequence being that children disadvantaged by illness or other life circumstances often bear greater risk than other children (Institute of Medicine [IOM] 2004, Nelson 2011, NHRPAC 2002, SACHRP 2005, Shah 2011). Defenders of the relative interpretation retort that under limited circumstances its use might be ethically defensible, insofar as this standard would allow for incremental improvements to the lives of children who are socially disadvantaged (Snyder, Miller, and Gray 2011).

Proponents of the routine examinations interpretation (e.g., Fisher, Kornetsky, and Prentice 2007; Kopelman 2004a; Resnik 2005) argue that daily life risks—even those of average, normal, healthy children—might sometimes be too high to be considered acceptable or are otherwise inappropriate as a normative standard for research risk. For example, Kopelman asserts that “even if most children play in traffic at one time or another, that is not a morally justifiable basis for judging when research risks are comparably minimal” (2004a, p. 362; see also; Thompson 1990). However, a routine examinations standard has been criticized as being too conservative a standard for acceptable risk (Wendler 2005; Westra et al. 2011, Binik 2014).

The “charitable participation” and “undue burden” standards, while intriguing, do not obviously get around the preceding difficulties. Perhaps charitable activities might sometimes present risks that are too high, e.g., because parents did not properly think through the child’s participation in that activity. Or, conversely, perhaps the charitable participation standard is too conservative, with risks encountered outside of this context still sometimes qualifying as “minimal.” Similarly, there is no guarantee that the daily life risks faced by children who are not “unduly burdened” will all be sufficiently low to be considered “minimal.”

Finally, some commentators have highlighted the fundamental need for normative justification for exposing children to particular levels of risk in non-beneficial research and have questioned the adequacy of the current regulations in providing such justification. For example, Westra and colleagues have asserted that the current definition of minimal risk “confuses the descriptive and the normative” by flatly asserting the acceptability of daily life and routine examination risks (Westra et al. 2011, p. 497). Kopelman (2004a, p. 361) and Binik (2014, p. 4) have also made similar claims.

Regarding the category “a minor increase over minimal risk,” little has been written. Wendler and Emanuel (2005) published an analysis where they considered five possible interpretations of this risk category, eventually settling on a “socially acceptable risk” interpretation defined as those risks that are greater than daily life risks but still considered socially acceptable. However, understood as a descriptive judgment, it might be argued that “socially acceptable” begs the question, and understood as a normative judgment, it is an open question how we should determine what is “socially acceptable.”

Two important features of these risk debates deserve note. First, most commentators have taken a stipulative approach to defining minimal risk, according to which minimal risk is defined as the set of risks associated with certain activities. This argumentative strategy attempts to solve a “Goldilocks problem” by identifying a clear definition of minimal risk that gets it “just right” without being overprotectionist or underprotectionist in nature. This point applies not just to discussion of daily life and routine examination risks, but also to newer proposed standards, such as the charitable participation and undue burden standards, since these also define minimal risk in terms of the risks associated with particular activities. Further, many commentators appear to tacitly assume that IRBs are rigidly bound to whatever definition of minimal risk is chosen, since the identification of a counterexample to a proposed interpretation (e.g., that it would allow for a particular risk thought to be excessive) is typically taken as a reason to reject the interpretation. This seems to presuppose little discretion in the application of the risk categories.

Second, intuition plays a large role in many discussions about the interpretation of the risk categories. By “intuition,” we mean what seems to be true to someone upon reflection. Quoting Resnik: “In layman’s terms an intuition is a gut feeling or hunch. Intuition is usually distinguished from reasoning, which involves forming beliefs or judgments as a result of conscious inference or deliberation” (Resnik 2017, p. 4). In many cases, a commentator’s rejection of one or more interpretations of minimal risk rests upon their intuitive judgment that the interpretation establishes an excessively low or high risk threshold. To take just one example, Binik (2014) writes that “the healthy child interpretation may not successfully restrict nontherapeutic research procedures to a sufficiently low degree of risk.... Children living in poor neighborhoods or in geographically dangerous areas face daily risks that *seem impermissibly high*” (p. 4, emphasis ours). Such intuitive judgments are widespread in the published literature on minimal risk and might partly explain why disagreement persists: for persons not inclined to agree with a given assertion about the nature of minimal risk, there is not much in the way of argument to persuade them. This intuitive character also raises questions about rational standards in risk evaluation: apart from intuition, how—if at all—might we go about justifying claims that a

particular risk is minimal or more than minimal, or that a research risk exceeds or falls below an accepted comparison standard?

Taking stock of the risk debates: Exactly what is being debated?

Pediatric research raises numerous and distinct ethical issues, and these are not always sufficiently distinguished in debates over the interpretation of minimal risk. Making progress in the risk debates, therefore, requires that we clarify the exact issue(s) being debated. Doing so will allow us to more clearly appreciate the role of intuition in these debates and the limitations of stipulative approaches to defining minimal risk.

We should first distinguish the question of which specific risks are acceptable in non-beneficial pediatric research from the more fundamental question of why such research might be justified in the first place. The National Commission's answer to the latter question appears to have been a "considered ethical judgment"—not something that is logically entailed by prior ethical principles or theory, but instead an "all things considered" judgment that seemed, upon much reflection, to be most ethically defensible to its members.⁴ The Commission considered and ultimately rejected the position that, because children cannot provide authentic informed consent to such research, it is never morally permissible to perform (see Jonsen 2006). Part of the Commission's reasoning in arriving at this conclusion was that the disallowance of all non-beneficial pediatric research would unduly hamper pediatric drug development. In this scenario, while some children might be spared from being exposed to risk in the research setting, such risk would be shifted onto other children who, for example, would be receiving medications whose safety and efficacy were inadequately assured.

However, the Commission's reasoning in arriving at this conclusion was not simply consequentialist. Rather, it stipulated that the ethical acceptability of such research depended on its being low risk—or, in the Commission's terminology, "minimal risk." As well, the Commission reasoned that many daily life risks arise in activities regarding which society grants parents discretion to involve their children. Since a child's daily life experiences involve some degree of risk, even under favorable circumstances, the Commission posited that children's enrollment in low-risk research also falls within the scope of appropriate parental decision-making.⁵

Neither the regulations nor *Research Involving Children* explain exactly why daily life risks and routine examination risks were chosen for the definition of "minimal risk," but it seems reasonable to conclude that the Commission wanted to clarify the nature of the risks that it considered minimal in magnitude; the primary normative significance of these activities seems to be that their associated risks were thought to be low. The regulatory definition of

⁴DeGrazia (2003) defines a considered judgment as follows: "I suggest that considered judgments, which have an initial credibility (though not infallibility), are moral judgments that are adequately informed in terms of both factual information and relevant moral alternatives, stable even when one is not under pressure to accept them, free from personal conflict of interest, and the like" (p. 221).

⁵To wit: "the scope of parental authority routinely covers a child's participation in many activities in which risk is more than minimal, and yet benefit is questionable" (Research Involving Children 1978, p. 2110). And: "The permission that parents give for children's participation in research can be accepted as an exercise of their general role, as caretakers, to guide decisions affecting their children's lives and activities" (Research Involving Children 1978, p. 2110).

minimal risk should be viewed as an attempt to answer the question of which specific risks (or risk equivalents) are acceptable in pediatric research not offering the prospect of direct benefit. This definition specifies and operationalizes the principle, “non-beneficial pediatric research is ethically acceptable when (inter alia) the risks are low.”

The pediatric regulations under 45 CFR 46 subpart D (and the corresponding FDA regulations at 21 CFR 50 subpart D) do not contain ethical discussion or reference *Research Involving Children*. However, knowing that the pediatric regulations derive from this report, we can appreciate that the National Commission did not think that risks should be considered acceptable in non-beneficial pediatric research “just because” they are encountered in daily life or routine examinations. Rather, the definition of minimal risk contained in the regulations is the product of ethical deliberation about the permissibility of non-beneficial pediatric research and its justificatory conditions. Therefore, while the regulatory definition of minimal risk might be vulnerable to other criticisms, it does not conflate the descriptive and the normative.⁶

From this departure point, one direction in which to take the conversation is to ask whether additional ethical justifications for non-beneficial pediatric research are possible. This is an important question that has received some recent attention (see, e.g., Wendler 2012), but it is also orthogonal to our present concern, which is the appropriate interpretation of “minimal risk.” Making progress in this debate requires that we stop and think about the kind of judgment that minimal risk is supposed to represent. We recognize two possibilities: 1) minimal risk is a judgment about risk magnitude, and 2) minimal risk is a judgment about which risks are acceptably low. These two understandings are related but not identical.

The term “minimal,” like the terms “small,” “low,” “moderate,” “high,” or “excessive,” is a qualifier of magnitude; thus, the most straightforward way to understand the term “minimal risk” is as a judgment of risk magnitude. This interpretation seems consistent with much of the National Commission’s discussion in *Research Involving Children* and with much subsequent commentary. In this view, controversy arises in the interpretation of minimal risk because some proposed interpretations (e.g., the relative interpretation of daily life risks) include risks that are not actually minimal in magnitude, but instead are of a higher magnitude. As well, controversy might arise because some proposed interpretations (e.g., the routine examinations standard) exclude risks that, while higher in magnitude, are nonetheless thought to be minimal. In both cases, the relevant problem is that the proposed interpretations misidentify the magnitude of certain risks.⁷

⁶Said another way, we should distinguish between a definition that is normatively inadequate and a definition that fails entirely to recognize the normative character of a term. The Commission’s discussion of minimal risk gives no indication they thought that just because something is the case (children are exposed to some risks in daily life) it therefore ought to be the case (they should be exposed to similar risks in research); that is not the logic of their discussion.

⁷One minor complication to this point is that the “relative versus absolute” debate seems to hinge primarily on considerations of justice, rather than considerations of risk magnitude. However, we believe that this is best viewed as a second-order issue in the interpretation of “minimal risk,” since this issue arose only because of ambiguity in the wording of the risk categories and does not concern their basic meaning. Further, the issue here is not simply the abstract one that different children would be exposed to different levels of risk absent a good reason. Rather, a significant part of the concern at issue is that some children might be exposed to risks that are *too high* to rightfully be considered “minimal” or “low”—that without additional qualification, “the risks of daily life” are not a good substitute for risks that are minimal in magnitude.

It might be argued that the type of judgment being made when enrolling a child into non-beneficial research is not simply that the associated risks are low, but rather that they are *acceptably* low (i.e., “low enough”). Such a judgment combines considerations of risk magnitude and normative acceptability. The main reason for understanding minimal risk in this way is that not all low risks might be viewed as acceptably low in the context of a particular pediatric research intervention not offering the prospect of direct benefit; an additional “all things considered” judgment might be required here.

Numerous commentators in the research ethics literature seem to understand minimal risk in this way. For example, Nelson and Ross (2005) argue for a “scrupulous parent” interpretation of the pediatric risk categories, according to which acceptable risks in pediatric research are those to which a scrupulous parent should intentionally expose their child (Nelson and Ross 2005). Thus, IRBs would need to put themselves in the position of a hypothetical “scrupulous parent” when making decisions about whether research protocols present minimal risk or a minor increase over minimal risk. As well, the National Human Research Protections Advisory Committee (NHRPAC) commented that “the minimal risk standard defines a permissible level of risk in research as the socially allowable risks, which parents generally permit their children to be exposed to in nonresearch situations” (NHRPAC 2002, as quoted in Kopelman 2004a, p. 364). The motivation for the charitable participation standard also seems to be that the risks of charitable activities are more normatively appropriate as an equivalence standard for research risks than are daily life risks in general.

On either view about the fundamental meaning of minimal risk, progress in resolving debates about its interpretation depends upon clarification of the rational standards that might apply to judgments about minimal risk. We turn to this issue in the next section, where we discuss the evaluative components of risk assessment and the limitations of current approaches to interpreting minimal risk.

Pediatric risk assessment: Limits of current approaches

In considering what rational standards might apply to judgments about minimal risk or its derivatives, it should first be emphasized that “risk” is a moral concept, signifying a possible harm and incorporating both the probability of the harm’s occurrence and the magnitude of the possible harm, should it occur. To harm someone is to make them worse off or set back their interests. Determining whether someone has been harmed thus requires that we consider how his or her welfare is affected by the action or event in question, and this consideration is by definition moral in nature (Resnik 2017; Rossi 2012).

One reason why this is important to clarify is that our judgments about the harmfulness of many things (e.g., that cancer is harmful) are likely to be uncontroversial and therefore overlooked. However, the harmfulness of some things might be a point of contention, and here it is important to be aware of the evaluative dimensions of such judgments so that people do not talk past each other. For example, suppose that an experimental pediatric medication raises the level of a serum biomarker of unknown significance. Should this be considered a harm? It is not obvious.

In addition, the evaluative status of judgments about risk can be overlooked because of the emphasis often placed upon probability assessment in branches of the natural sciences dealing with risk, for example in the field of environmental risk analysis. Because probability assessment incorporates complex and often quantitative scientific methods, this has historically led some actors to assert that risk assessment is a value-free process (see Shrader-Frechette 1991). A similar problem might arise in the context of human-subjects research, where multiple studies have documented disagreement and inconsistency among Institutional Review Boards (IRBs) and IRB chairs as to whether particular procedures (e.g., a spinal tap) should be considered “minimal risk” or a “minor increase over minimal risk” (Janofsky and Starfield 1981; Shah et al. 2004; Wendler and Varma 2006). There are numerous possible explanations for this variability, but one of them is that IRBs often operate without robust empirical information about the nature and probability of harms posed by research interventions or comparator activities. In response to the preceding problem, some scholars have undertaken the task of gathering empirical information about the risks that individuals face in daily life or from research interventions (Wendler et al. 2005, Rid and Wendler 2011b), and this focus might give the impression that improving research risk assessment is primarily an empirical matter. Such empirical data can undoubtedly aid in applying the current federal regulations, but we should not expect more comprehensive empirical information to fully settle debates about risk magnitude.

In addition to the evaluative judgment that a particular outcome is harmful, judgments about risk magnitude seem also to involve two additional value judgments: first, a value judgment about the magnitude of the harm, and second, an “all things considered” value judgment about the magnitude of the risk, which incorporates both the magnitude of the harm and its probability of occurrence. Each of these three value judgments seems conceptually distinct: two persons might agree that something (e.g., breaking a leg) is harmful, but disagree about how harmful it is. Similarly, two persons might agree in their evaluation of the magnitude of a given harm, as well the probability of the harm occurring, but still disagree about the magnitude of the risk in question (Resnik 2017; Rossi 2012).

To arbitrate disagreement when it arises, something must be said about how such judgments *should* be made. The dominant strategy for interpreting minimal risk and its derivatives does not directly address this issue, but instead stipulates that these terms should be defined as some set of risks. However, there is precedent for this kind of argumentative strategy in ethics, and it is vulnerable to well-known criticisms.

A major historical preoccupation of moral philosophers has been defining the nature of “good.” Philosophers have been interested not merely in the question of which specific things are good (or bad), but also in the more fundamental question of *what makes them good (or bad)*. One possible answer to this question, favored by some utilitarian philosophers, is hedonism: good things are good because they are pleasurable, and bad things are bad because they cause pain or deprive us of pleasure. In response to this assertion, British philosopher G.E. Moore elaborated his now-famous “open-question argument” (Moore 1902/1988), and Moore’s arguments are as relevant to contemporary discussions about research risk as they were to utilitarianism a century ago.

Moore's focus was on the meaning of moral language. He argued that if "goodness" simply meant "pleasure," then we should be able to substitute the latter word for the former without changing the meaning of a sentence. Thus, the sentence "Is pleasure good?" would have the same meaning as the sentence "Is pleasure pleasure?" Moore argued that these two sentences do not in fact mean the same thing: the former question is "open" in the sense that it is both intelligible as a question and open to doubt or debate. However, the latter question is "closed": it is not open to doubt or debate, and it is barely intelligible as a question; by definition, pleasure is pleasure. This "open question argument" can be applied to any other attempt to define goodness in terms of a particular property (e.g., happiness, what we desire). For Moore, the failure of other terms to adequately substitute for "good" in linguistic usage meant that goodness could not be *defined* in terms of any single property, even though the properties in question (happiness, pleasure, etc.) might be *examples of* good things (Moore 1988).

The same argument can be applied to the definition of "minimal risk." If, for example, "minimal risk" just means "the risks of daily life," then we should be able to substitute the term "risks of daily life" for the term "minimal risk" without any linguistic problem. Hence the statement "I know risk *x* is encountered in daily life, but is it minimal?" would have an equivalent meaning to "I know that risk *x* is encountered in daily life, but is it encountered in daily life?" However, the latter statement does not seem to have the same meaning as the former: to paraphrase Moore, we seem to have a different thing before our mind when we ask whether such risks are minimal. Thus, any proposed definition of "minimal risk" in terms of a specific set of risks fails the open-question test.

Philosophical work since Moore has recognized the possibility that a proposed definition could be valid despite failing the open-question test. For example, some philosophers have argued that "water" and "H₂O" do not mean the same thing linguistically, even though they *are* the same thing. The best response to this rejoinder is to see the open-question argument as shifting the burden of proof (see Strandberg 2004). The failure of a definition to pass the open-question test means that proponents of the definition must provide strong arguments to show that two terms are identical, even though they do not seem to mean the same thing. In the case of water, we have considerable scientific theory and evidence supporting the claim that the chemical formula of water is "H₂O" and thus that the two terms are in fact identical. However, there are no such arguments that appear to support the definition of "minimal risk" (or its derivative terms) as a set of risks. Any particular risk might be an *example of* a "minimal risk," but this is a different matter than saying that "minimal risk" can be adequately *defined* by a particular set of risks.

The best we can hope for is that a particular set of risks will adequately capture people's intuitions about what constitutes minimal risk, but as we have already seen, all such definitions are at odds with the intuitions of at least some persons.

If "minimal risk" cannot be adequately defined by any set of risks, are there perhaps other criteria that apply to making judgments about risk magnitude in a rationally defensible way, such that we can arbitrate disagreement when it arises? Resnik (2017) has recently reviewed the role of intuition in risk-benefit assessment in research and has argued that all ethical

judgments involve at least some intuitive components. However, he argues that some judgments based on intuition might be “replaced by another belief or judgment obtained by cogent reasoning” (p. 10). As concerns moral judgments, Resnik argues that they are replaceable to some degree, for example “as a result of obtaining more information, engaging in further reflection, or both” (p. 11). This certainly seems correct: few if any philosophers would argue that ethical judgments reduce entirely to intuition. Aside from the fact that ethical judgments often depend on truth-conditional empirical premises (e.g., relating to the effects of an action or policy), ethical justifications for conduct (e.g., exposing someone to risk or harm) often appeal to ethical rules, principles, and criteria. As well, ethical judgments depend on standards of argumentation derived from logic (e.g., entailment, refutation by counterexample, consistency, the avoidance of certain fallacies, etc.). Through the application of these argumentative tools, we hope to arrive at defensible conclusions about what we ought to believe and what we ought to do. We also hope to change others’ minds by showing that they must, *on pain of irrationality*, accept a particular conclusion given certain agreed-upon premises and standards of logic.

However, the degree to which ethical intuitions are replaceable would seem to depend on the kind of ethical judgment in question. Philosophers conventionally distinguish between judgments about prudential value (what is good and bad for us) and judgments about ethical conduct (what we should and should not do)—or, to use more traditional language, between “the good” and “the right.” Judgments about risk magnitude fall into the former category, insofar as the concept of risk is derivative of the concept of harm and signifies something that might set back our interests. Many of the preceding standards of rationality apply to judgments about ethical conduct, but would seem to apply less to judgments about what is beneficial or harmful to us. As Resnik points out, “judgments or beliefs concerning personal preferences, tastes, pleasure, pain, discomfort, or offensiveness” would seem “not judicable by reason” (p. 8). He offers as an example two persons arguing over whether a blood draw or headache is more painful—here, “reasoning would seem to have little power to convince them one way or the other” (p. 8).

Resnik is not alone in asserting this: the idea that judgments about “the good” are less amenable to rational arbitration than judgments about “the right” is a longstanding one in philosophy, argued by contemporary theorists (see, e.g., Crisp 2013; Heathwood 2006) and such historical luminaries as Moore (1902/1988) and John Stuart Mill (1861/2004). Moreover, the role of intuition seems magnified when we consider the *degree to which* something is beneficial, harmful, or risky. We can appeal to shared preferences or values (e.g., a shared judgment that pain or loss of function is bad) to show why something is beneficial or harmful, although if these preferences or values are disputed then rational adjudication might not be possible. But once two persons are agreed on the fact that something is beneficial, harmful, or risky, the weighting of the magnitude of such benefit, harm, or risk would seem heavily intuitive (see, e.g., Feinberg 1984, p. 203; Resnik 2017, pp. 17–8, Rossi 2012).

The preceding comments apply to the interpretation of minimal risk as a judgment of risk magnitude, but they also apply to the “acceptably low risk” interpretation of minimal risk. For starters, for a risk to be *acceptably* low, it must first be low; we presume that a

scrupulous parent would not enroll a child in research not offering the prospect of direct benefit if the associated risks were significant. But this just returns us to the philosophical issues about risk magnitude already considered. On this point, we can observe that the reason why certain activities, such as charitable activities, might be thought more normatively appropriate than other daily life activities as a comparator for research risks is that their risks are more likely to be minimal in magnitude. Some daily life risks are not run voluntarily because they are low in magnitude, but instead are run because they are part of an activity that is unavoidable or not reasonably avoidable, in which case there is no reason to expect that such risks will be minimal in magnitude. Similarly, some daily life risks are run because they are part of an activity offering compensating benefit, and where such benefit exists a person might tolerate higher risk (analogously to pediatric research offering the prospect of direct benefit). However, charitable activities are both optional in nature and without compensating benefit, and if a parent is acting scrupulously then we might expect that they would only allow their child to participate in such activities when the risks are low. Even so, as already discussed, it seems unlikely that charitable activity risks will perfectly capture our intuitive judgments about what constitutes minimal risk.

Second, any attempt to define “acceptably low risk” in terms of a set of risks runs into the same open-question argument discussed above; the philosophical problems are the same. Third, the scope of ethical principles is sometimes “fuzzy,” meaning that reasoning can only get us so far in justification, with intuition playing an inextricable role in an individual’s application of the principle. Non-beneficial pediatric research seems a good candidate for such a case: once we arrive at the principle that such research is ethically permissible only when the risks are kept acceptably low (and other criteria are satisfied), it seems dubious that any principled justification could be provided for why certain risks are “acceptably low” while others are not. Judgments about acceptably low risk might appeal to a number of normative considerations—for example, the nature of the research question being asked, its perceived importance, its likelihood of leading to medical progress, the child’s personal connection to the condition being studied, and how s/he as an individual would react to the study interventions being performed⁸—but a decision at this level of specificity seems highly contextual and not amenable to being captured by a more specific principle or set of criteria. Certainly, no one has yet been able to articulate what these might be.

Finally, it should be noted that evaluative intuitions also seem to loom large when we turn to the issue of risk comparison. Here the problem is that the risks of research interventions are often qualitatively dissimilar to the risks of daily life or routine examinations, and furthermore that both research interventions and comparator activities might involve a range of different risks, each with different probabilities and harm magnitudes. Thus, it can be unclear whether a risk associated with a research intervention is or is not greater than another risk to which it is being compared. A framework has been proposed for systematically comparing research risks to an established standard, which would involve

⁸These latter features are already supposed to be taken into account when IRBs consider whether a non-minimal-risk intervention might qualify as a minor increase over minimal risk. For minimal risk studies, IRBs might think about whether research risks are “reasonably commensurate” (45 CFR 46.406) with a child’s prior experiences, at least if a relative interpretation of minimal risk is ever appropriate. As well, once a study receives IRB approval and is recruiting participants, parents would need to think about how *their individual child* would fare in the research when making an enrollment decision.

ranking harms on a 5-level scale and then comparing the probabilities of these harms (using information that might be obtained from more rigorous efforts to empirically characterize comparator risks) (Rid, Emanuel, and Wendler 2010). However, this procedure involves numerous contestable value judgments and so presents a controversial solution to the philosophical problem of risk comparison. For example, this framework precludes comparison of risks across harm categories, even though a low probability of a moderate harm might be prudentially equivalent to a higher probability of a small harm (Rossi and Nelson 2012).

Thus, to return to the central issue, judgments about what constitutes minimal risk or a minor increase over minimal risk would seem to be heavily intuitive in nature, where intuitions are not highly “replaceable.” It is understandable that many stakeholders would desire an objective definition of minimal risk that leaves little interpretive room, but from a philosophical standpoint this seems overreaching.

On the interpretation of minimal risk and its derivatives: Can we move beyond intuition?

How, then, should we proceed with regard to the interpretation of the pediatric risk categories? In our view, we have two general options.

The first option is to choose one of the more concrete standards on offer (e.g., the routine examinations, charitable participation, or absolute daily life standard), understanding that what these standards sacrifice in philosophical defensibility they might make up for in other respects. Such standards may provide IRBs with some guidance, and they provide transparency and accountability to the public. Even with this approach there will be uncertainty in pediatric risk evaluation, since IRBs often operate with limited empirical information and since the process of risk comparison is indeterminate and value-laden. However, at least in theory the acceptable risk standard could be defined with some degree of concreteness.

On this approach, it would need to be acknowledged and accepted that a perfect, counterexample-free definition of minimal risk is not in the offing—meaning that bioethicists should stop trying to find one. As well, when picking which interpretation to use, we would need to revisit the fundamental issue of balancing scientific progress with child protection. Clearly, it would not do to have regulations that routinely allow children to be exposed to excessive risk. At the same time, an overprotectionist approach (e.g., entirely excluding children from non-beneficial research) would not do, since this would result, for example, in a lack of safe and effective medicines for children. Since a perfect balance between over- and underprotectionism is unlikely to be achieved through stipulative definitions of minimal risk, the problem would need to be reframed from a philosophical problem (choosing the definition of “minimal risk” that is objectively correct) to a practical problem (minimizing the kinds of overprotectionist or underprotectionist outcomes we wish to avoid).

We might draw an analogy here to type-I and type-II errors in statistics: just as we wish to avoid false positives and false negatives in statistical inference, so too do we wish to avoid the errors of exposing children to excessive risks or shielding them from reasonable risks and setting back research in the process. And just as it is impossible in statistics to simultaneously minimize both type-I and type-II errors, it will be impossible to simultaneously minimize both kinds of errors in pediatric research: whatever definition of minimal risk is chosen might lead to one outcome or the other. In choosing how to interpret and apply the pediatric risk categories to a given protocol, we must think about which kind of error it is worse to make and choose the interpretation that minimizes this. In choosing, we will also have to think about the relative importance of consistency and ease of application in IRB judgments. Skeptics might argue that this approach merely trades one problem for others, particularly since intuition will still be involved in these determinations. However, if we stop trying to achieve perfection and instead adopt a consensus-driven approach aimed at identifying the best available compromise, resolution of the risk debates might be possible.

A second approach is to accept the intuitive nature of judgments about minimal risk and its derivatives and to build an alternative model of risk evaluation that helps to improve the quality of such judgments. Even though judgments about minimal risk do not seem amenable to rational arbitration by specific definitions, it still seems incorrect to say that one person's intuitive judgments about what constitutes minimal risk will always be equally plausible to another's. For example, a person's judgment about risk magnitude seems more reliable with accurate information about the probability of a harm's occurrence than without it. Making ethical judgments in a defensible manner need not be construed only as a function of justification by valid rules or principles. Objectivity can also be construed as the result of making ethical judgments under conditions thought to improve the reliability of such judgments.

One family of approaches in this tradition is referred to variably as "dispositional," "ideal spectator," or "ideal observer" theories (see, e.g., Firth 1952; Harman 2000; Kawall 2006, Smith, Lewis, and Johnston 1989). Dispositional models of ethical judgment are "response-dependent": they posit that ethically correct judgments are constituted by the hypothetical responses of an ideally situated judge (a.k.a. "spectator" or "observer"). For example, on this approach the question of whether murder is wrong would be answered by appealing to whether an ideal spectator/observer would disapprove of it. Similarly, the question of whether a risk is minimal or greater than minimal would, on a dispositional account, be answered by appealing to whether an ideally situated person would judge it to be so.

Applying a dispositional theory to pediatric research risk evaluation requires that we elaborate a list of conditions thought to improve the quality of such judgments, which again are understood here as pertaining to "low" or "acceptably low" risk. As a preliminary list, we suggest the following:

- Having the most accurate empirical information about the nature and likelihood of harms faced by children in daily life, routine examinations, charitable

activities, or other comparator activities, as well as those posed by specific research interventions.

- Being able to accurately imagine or understand what it would be like to experience a particular harm—what philosophers have variously called “omnipercipience” or “imaginative acquaintance” (see Firth 1952, Smith, Lewis, and Johnston 1989). As concerns pediatric research, it is particularly important to consider that certain harms might affect children differently than adults.
- Explicitly considering information about features relating to the evaluation of harm magnitude, such as the duration and reversibility of the harm (see Bogardus et al. 1999; Rid, Emanuel, and Wendler 2010).
- Being able to compare evaluations of different risks, both by one’s self and others, with respect to their magnitude and their normative acceptability in specific contexts. This would allow for better calibration of judgments about risk and checks for consistency or inconsistency in judgment.
- Making evaluations under conditions free of distorting bias, e.g., from personal conflict of interest. For example, because of so-called “affiliation bias,” scientists may judge their peers’ work to be less risky and of greater importance than nonscientists would judge (see Kimmelman 2004).
- Having good information about the context for a research study, the importance of the research question being asked, and its likelihood of resulting in medical advancement.

Additional conditions might of course be added. Such conditions would have to be the result not only of philosophical argument, but also of group deliberation. For the model to gain practical traction, the conditions of good ethical judgment must be conditions that are broadly agreed to and not merely asserted by one or a few parties. This suggests that, in the longer term, open and public collaborative efforts (e.g., workshops) within the research ethics community to identify and flesh out such conditions, including broad representation from all those affected by the deliberations, would be a good idea. In the short to medium term, the conditions we have listed seem like a good start, and attention would then need to be paid to the structuring of IRB review to promote reflection under these conditions. In part, this would require that IRBs have as much relevant information available to them as is possible. Thus, we support prior suggestions to better empirically characterize the risks of research interventions and comparator activities and to establish a research risk repository to facilitate IRBs’ access to this information (Wendler et al. 2005, Rid and Wendler 2011b). However, having access to relevant information is not just a function of empirical research; it is also a function of IRB representation and procedures. This brings us to a second philosophical tradition upon which we might draw when attempting to improve upon sheer intuition in research risk assessment: deliberative democracy.

Deliberative democratic theories emphasize that defensible public policies are the result of a reason-giving, inclusive, and mutually respectful process of deliberation whereby all parties are able to express a viewpoint and are expected to both advance reasons for their view and to respond to reasons and criticisms advanced by others (Gutmann and Thompson 2004).

Philosophical motivation for these theories comes from multiple sources, including counterexamples to simple majoritarianism (e.g., majority vote leading to policies that clearly seem morally wrong), the requirement that citizens be treated as autonomous agents and not merely as governed subjects, and the belief that at a fundamental level, defensible laws and policies must be justified through appeal to public reason (Gutmann and Thompson 2004).

Deliberative democratic theories are useful to draw upon in the IRB setting for several reasons. First, human-subjects research is an essentially public activity: it is of significant public interest given its importance to medical progress and the ethical concerns implicit in its conduct; is subject to public regulation; and often is paid for using public tax dollars. The ethical justification of such research therefore involves not just first-order considerations about its actual conduct, but also second-order (i.e., political) considerations pertaining to its transparency, regulation, and public justification.

Second, attention to the procedural aspects of IRB deliberation about risk can serve as a way of implementing the dispositional model of risk evaluation introduced above. For example, efforts to further diversify IRB representation can help to mitigate potential bias (see Kimmelman 2004) and to increase the “imaginative acquaintance” of the deliberating body, for example by including individuals who have first-hand familiarity with unusual harms of concern or who can articulate the experience of research participants. In the context of pediatric research, this might include the perspectives of parents whose children have been previously enrolled in research studies of a similar nature to that under review, and perhaps also of older children (e.g., adolescents) who themselves have been involved in similar research.

Attention to procedure can also help to ensure that risk evaluations will be made in a careful and explicit manner, per the reason-giving requirement of deliberative democracy. Explicit justification does not have to mean “justification beyond a reasonable doubt”; it simply means justification that is as clear and cogent as is possible to give. This end might be aided, for example, by using harm-ranking scales to guide judgment (see Rid, Emanuel, and Wendler 2010) or step-wise procedural algorithms to help structure risk evaluations. Here it just needs to be kept in mind that decisions resulting from the use of such aids are inherently value-laden. One possibility for structuring the risk evaluation process is as follows:

- a.** Consider the probability of relevant harms and the strength of evidence for them. Attempt to address and arbitrate any disagreement here.
- b.** Consider the magnitude of the potential harms, possibly using an agreed-upon harm ranking scale. Attempt to address and arbitrate any disagreement here.
- c.** Intuitively evaluate the overall risk level based on probability and harm magnitude.
- d.** Compare this intuitive risk judgment to other risks that the rater deems of similar magnitude, as well as examples of risks from well-defined categories (daily life, routine examinations) that are good candidates for “minimal” risks.

- e. Address discrepancies/inconsistencies in risk evaluations between the risk under consideration and comparator risks. Address the adequacy of comparator risks as being “minimal” (e.g., is the daily life risk being compared to the research risk really “minimal”?). Address inter-rater discrepancies. Seek out possible sources of bias.
- f. Attempt to resolve any remaining disagreement between risk evaluators. If disagreement persists, implement a procedure for resolving disagreement (see below).

Many of our suggestions here are consonant with Resnik’s (2017), who builds upon a preexisting framework for risk-benefit assessment (Rid et al. 2011a) to identify places where intuition is likely to factor into such judgments and might be reduced. Regardless of which type of procedural aids to risk evaluation are used, we can be reasonably confident that the more robust deliberation is, the more likely IRBs will be to uncover relevant information, reflect thoughtfully about risk, and arrive at decisions that are well-justified. This process contrasts with a model in which decisions about acceptable risk are made simply by majority vote in an aggregative fashion, which “takes ... preferences as given” and “requires no justification for the preferences themselves” (Gutmann and Thompson 2004, p. 13).

While IRBs are currently required to record minutes of committee meetings, some commentary suggests that such minutes are often insufficiently detailed (e.g., US FDA 2015). A well-implemented deliberative approach would include IRB meeting minutes that are detailed enough to provide a clear justification of the rationale for a decision and any controversy or disagreement. As well, the extent to which IRB meeting minutes are used to guide future meetings—as opposed to being filed away—is unclear, but such records of justification could provide a helpful resource for future deliberation by the committee. If suitably anonymized, IRB deliberations about particular research interventions or comparator risks could also be made a publicly available resource, analogously to a proposed research risk repository (Rid and Wendler 2011b).

Third, given their heavy intuitive components, it seems reasonable to posit that risk evaluations will be somewhat pluralistic in nature, with different persons coming to different, equally defensible conclusions even under favorable conditions of reflection. When a single, correct conclusion is beyond the reach of rational argument or when consensus cannot be obtained about the most appropriate course of action, ensuring that decisions are made according to defensible procedures can help to legitimize them. Along with this, discursive justification takes on heightened significance when guiding principles are fuzzy in their application, as they seem to be in making decisions about acceptably low risk in pediatric research. Even though judgments about acceptably low risk seem highly contextual, such decisions seem better justified when they are the result of robust reflection and deliberation as compared to when they are the result of pre-reflective intuition.

We should emphasize that our proposed approach does not prevent IRBs from considering daily life, routine examination, or charitable activity risks as possible examples of minimal risk and using such risks as a baseline for deliberation. However, some flexibility should be retained when employing such comparison standards, given the philosophical problems

attending to their rigid use; IRBs should have some “deliberative space” to consider whether a risk sanctioned by a particular standard really should be regarded as minimal.

To what degree would a dispositional, deliberative approach provide for objectivity in judgments about risk magnitude? This is an open question. To settle it, we would need to elaborate a fuller list of features to which deliberating bodies (e.g., IRBs and other research ethics committees) should attend, develop procedural models for incorporating such features into research protocol review, and conduct empirical assessments to determine the degree of convergence in opinion that such methods provide. While some philosophers writing in the dispositional tradition have assumed that it is absolutist (i.e., capable of identifying indisputably correct answers), we make no such assumption here. As concerns the identification of ideal conditions in the dispositional model, the most that we may be able to rely on is our considered judgments about which conditions promote defensible judgments, and some degree of pluralism might be expected here. As well, even when ideal conditions are agreed-upon and individuals’ evaluations conform to them, different individuals might still produce different evaluations. At this point, the settling of disagreement would not be based on arbitration by a rational standard, but instead on procedural justification.⁹ A dispositional, deliberative model of risk evaluation does not demand that all parties agree in their evaluations before a decision is made. IRB deliberation cannot continue indefinitely and at some point a decision must be made. Our proposed approach does not preclude the use of majority vote to decide protocols when IRB members do not reach consensus, but instead emphasizes the nature of the deliberative process leading up to this vote.¹⁰

Further, pluralism in risk evaluation is not a fatal concern. The attraction of our dispositional, deliberative approach is not that it offers an alternative way to arrive at absolute objectivity in risk evaluation—something that we think is not possible—but instead that it offers the best realistic prospects for improving upon judgments that by their nature are heavily intuitive. Empirical studies of IRB decision-making have found that within IRBs, deliberation in the face of disagreement can lead to consensus and closure (see Stark 2012). As well, in a test run of a risk comparison algorithm, Rid and colleagues (2010) note that deliberation was often able to resolve disagreement about the ranking of harm magnitudes. These findings provide reason to be optimistic that IRBs would be able to adopt a dispositional, deliberative model and not be deadlocked by disagreement. If anything, emphasizing a more structured deliberative process, one that is explicitly informed by both empirical data and specific moral considerations, would seem likely to improve upon the status quo in this regard. Inconsistent judgments *between* different IRBs might arise on this

⁹The model we propose shows some broad similarities to the coherence model of ethical justification, often described as “reflective equilibrium.” Classically, the equilibrium is between our judgments about the ethical principles that should guide conduct and our judgments about what to do in specific cases: each informs and can revise the other. Our proposed approach is not principle-driven: we do not think that defensible judgments about minimal risk can be modeled as the output of specific ethical principles. Thus, this type of equilibrium would not apply. However, coherence can be thought of more broadly as involving ethical judgments that are stable under critical reflection, well-informed by empirical data, free of bias, and consistent with other ethical beliefs. These are also features of the dispositional model, and as already discussed the ideal conditions identified by the dispositional model might be justified only by appeal to reflective judgment. Further, when thinking about minimal risk as a judgment of acceptably low risk, the normative component of such a judgment (i.e., the “acceptable” part) would seem to rest principally upon reflective intuition, since as argued earlier principle-based justification does not seem possible at this level of specificity.

¹⁰To wit: “Deliberation must end in a decision, but deliberative democracy does not itself specify a single procedure for reaching a final decision. It must rely on other procedures, most notably voting, which in themselves are not deliberative” (Gutmann and Thompson 2004, p. 18).

model, but so too might they arise under the status quo. We consider that issue in the next section.

Responding to concerns

Given the preceding sketch of a dispositional, deliberative model of risk evaluation, one important question concerns the degree to which this approach differs from what IRBs already do and from what other guidance has been suggested in the research ethics literature. Is it not already accepted that IRB decisions arise out of deliberation?

Here it is helpful to keep in mind that determinations about minimal risk or a minor increase over minimal risk are only one of several things with which IRBs are tasked. Additional IRB responsibilities include ensuring appropriate consent procedures, fair subject selection, and an appropriate balance between risk and benefit or knowledge. Fulfilling each of these responsibilities will sometimes require ethics committee deliberation, and this fact seems to be acknowledged on a conventional understanding of what IRBs do.

However, when we turn to the specific issue of interpreting the pediatric risk categories, reflection and deliberation have not been emphasized as much. Most guidance to date has focused on the elaboration of clear and stipulative definitions of minimal risk and its derivatives, with some commentators expressing concern about definitions that are procedural or liable to interpretation (Fisher, Kornetsky, and Prentice 2007; Wendler 2005; Wendler and Emanuel 2005).¹¹ As well, suggestions to “systematize” research risk assessment and comparison (Rid, Emanuel, and Wendler 2010; Wendler et al. 2005) sometimes seem to assume that IRB deliberation about minimal risk is a conditional phenomenon—that given enough empirical information and a systematic way to compare research risks to an established standard, IRB risk assessments could become exact and precise. In contrast, we believe that risk assessment is inherently value-laden and heavily intuitive and that an accepted model for IRB judgments should reflect this. Moreover, we should recognize a distinction between what IRBs actually do and what they would do on an idealized account. Both authors have served as members of IRBs, as well as other ethics committees (e.g., IACUCs), and can attest that committee decisions as they are actually made may not always live up to a deliberative ideal, even when it is acknowledged. As long as there is a gap between the actual and the ideal, there is reason to reaffirm that ideal.

A second concern has to do with the specifics of implementing our approach: Would it be too demanding, and would it be compatible with existing IRB procedures? Regarding full committee review, IRB deliberation about the pediatric risk categories would arise when protocols present greater than minimal risk or when their minimal-risk status is in doubt (e.g., the protocol was assigned under expedited review but deferred for committee discussion). Here, the changes we suggest seem to be ones of degree and not kind. IRBs are already expected to deliberate about any number of ethical considerations; our arguments simply emphasize that the pediatric risk categories are one of them. Existing regulations

¹¹One notable exception is Kimmelman (2004), who recommends a renewed focus on IRB composition and procedure so as to make IRBs more “nonarbitrary and rigorous, but also inclusive and comprehensive of different views of risk” (Kimmelman 2004, p. 387). This suggestion is entirely in line with our philosophical arguments here.

already attempt to ensure IRB diversity and allow for IRBs to invite outside subject-matter experts to committee deliberations. The model we propose suggests that additional measures should be taken to ensure diversity of representation and expertise on IRBs, but accommodating this suggestion should not require major structural changes.

As concerns the issue of time, the development of procedural algorithms to guide and inform risk evaluations would not, in our view, be likely to lengthen committee meetings and might actually shorten them, insofar as discussions would presumably be more focused. As well, deliberation about the risks of specific research procedures could produce “benchmark” judgments that could carry forward to future deliberations by the same IRB (either through institutional memory or a more formalized mechanism) or, if publicly available, to other IRBs as well. This could both economize deliberation—ensuring that the same ground is not continually re-tread—and improve the overall quality of risk evaluations.

One might worry that our proposed dispositional, deliberative model is incompatible with the expedited review category, insofar as expedited review relies on predetermined, well-defined categories of minimal risk research and is typically conducted by one reviewer out of committee. However, endorsing a deliberative model of ethical justification does not require deliberation in all circumstances: “deliberative democracy makes room for many other forms of decision-making (including bargaining among groups, and secret operations ordered by executives), as long as the use of these forms themselves is justified at some point in a deliberative process” (Gutmann and Thompson 2004, p. 3). Thus, if the expedited review categories are themselves arrived at in an ethically defensible fashion, their use once established is not prohibited in our model.

Similarly, review by a single IRB member is not prohibited in our model. Again, the goal of our proposed model is to improve upon the intuitive nature of risk assessments in the various contexts in which they occur. For expedited review conducted by single individuals, focus would be placed on making reflection about risk as robust as possible, with such reflection ideally conforming to conditions identified in the dispositional model. Empirical information and decision algorithms used in full committee review could presumably be used here too. Expedited review under the status quo already involves individual reflection and judgment, insofar as it is usually done without robust empirical information and insofar as risk comparison is uncertain and evaluative. Our approach emphasizes that the need for this reflection and judgment arises not from a deficiency in the expedited review process (e.g., an insufficiently systematic risk comparison process) but from the inherent nature of risk assessment.

Finally, while our proposed model is likely to result in some variability in judgment between different IRBs, we do not see this as a major ethical concern. From a philosophical standpoint, we should keep in mind that consistency in judgment between different IRBs is an ethical desideratum only to the extent that it can be shown that IRBs should arrive at some particular judgment. In the absence of a clear and rationally binding standard to govern what should be considered minimal risk, it is not obviously unethical for one IRB to decide that a particular study is minimal risk while another does not. From a practical standpoint, under the status quo it is already true that different IRBs come to different judgments, and

we think it unlikely that our model will produce more variability than this. Further, pending changes to the *Common Rule* require that multi-site studies use a single IRB (Federal Policy for the Protection of Human Subjects 2017), and once this changes goes into effect concerns about inconsistency will be somewhat less of an issue.¹²

To conclude, we do not view our proposed approach as requiring major amendments to the status quo of IRB risk assessment, in part because this status quo has always been somewhat at odds with the rhetoric of much guidance on how IRBs should operate. Determinations about when a research risk is minimal or a minor increase over minimal are by necessity evaluative, intuitive, and subject to change, and there has never been a time in which IRB judgments about these things have been exact and precise. Moving forward in debates about the interpretation of minimal risk requires that this reality be acknowledged and that we shift our focus from identifying a problem-free definition of minimal risk to elaborating and then implementing conditions to improve upon use of intuition in risk assessment. We hope that our arguments here provide a first step in this direction.

References

- Binik A. On the minimal risk threshold with children. *American Journal of Bioethics*. 2014; 14(9):3–12. DOI: 10.1080/15265161.2014.935879
- Bogardus ST, Holmboe E, Jekel JF. Perils, pitfalls and possibilities in talking about medical risk. *JAMA : the Journal of the American Medical Association*. 1999; 281(11):1037–41. DOI: 10.1001/jama.281.11.1037 [PubMed: 10086441]
- Crisp, R. Well-Being. *The Stanford Encyclopedia of Philosophy*. 2013. Accessed <https://plato.stanford.edu/entries/well-being/>
- DeGrazia D. Common morality, coherence, and the principles of biomedical ethics. *Kennedy Institute of Ethics Journal*. 2003; 13(3):219–30. DOI: 10.1353/ken.2003.0020 [PubMed: 14577458]
- Feinberg, J. *Harm to others: The moral limits of the criminal law*. New York, NY: Oxford University Press; 1984. p. 203
- Federal Policy for the Protection of Human Subjects. 82 Federal Register 12. 2017 Jan 19. <https://www.federalregister.gov/documents/2017/01/19/2017-01058/federal-policy-for-the-protection-of-human-subjects>
- Firth R. Ethical absolutism and the ideal observer. *Philosophy and Phenomenological Research*. 1952; 12:317–45. DOI: 10.2307/2103988
- Fisher CB, Kornetsky SZ, Prentice ED. Determining risk in pediatric research with no prospect of direct benefit: Time for a national consensus on the interpretation of federal regulations. *American Journal of Bioethics*. 2007; 7(3):5–10. DOI: 10.1080/15265160601171572
- Freedman B, Fuks A, Weijer C. In loco parentis: Minimal risk as an ethical threshold for research upon children. *Hastings Center Report*. 1993; 23(2):13–19. DOI: 10.2307/3562813
- Gutmann, A., Thompson, D. *Why deliberative democracy?*. Princeton, NJ: Princeton University Press; 2004.
- Harman, G. *Explaining value and other essays in moral philosophy*. New York, NY: Oxford University Press; 2000. Moral agent and impartial spectator; p. 181-95.
- Heathwood C. Desire satisfactionism and hedonism. *Philosophical Studies*. 2006; 128:539–63. DOI: 10.1007/s11098-004-7817-y
- Institute of Medicine (IOM). *The ethical conduct of research involving children*. Washington, D.C: National Academies Press; 2004. p. 4

¹²The increasing reliance on regional IRBs makes it all the more important that the deliberative process be robust and transparent, insofar as a greater number of parties will be affected by the decisions of a single committee.

- Janofsky J, Starfield B. Assessment of risk in research on children. *Journal of Pediatrics*. 1981; 98:842–46. DOI: 10.1016/S0022-3476(81)80865-7 [PubMed: 7229775]
- Jonsen A. Nontherapeutic research with children: The Ramsey versus McCormick debate. *The Journal of Pediatrics*. 2006; 149:S12–S14. DOI: 10.1016/j.jpeds.2006.04.044 [PubMed: 16829235]
- Kawall J. On the moral epistemology of ideal observer theories. *Ethical Theory and Moral Practice*. 2006; 9:359–74. DOI: 10.1007/s10677-006-9016-8
- Kimmelman J. Valuing risk. *Kennedy Institute of Ethics Journal*. 2004; 14(4):369–93. DOI: 10.1353/ken.2004.0041 [PubMed: 15812985]
- Kopelman L. Minimal risk as an international ethical standard in research. *Journal of Medicine and Philosophy*. 2004a; 29(3):351–78. DOI: 10.1080/03605310490500545 [PubMed: 15512977]
- Kopelman LM. What conditions justify risky nontherapeutic or ‘no benefit’ pediatric studies: A sliding scale analysis. *Journal of Law, Medicine and Ethics*. 2004b; 32(4):749–58. DOI: 10.1111/j.1748-720X.2004.tb01980.x
- Mill, JS. Utilitarianism. In: Ryan, A., editor. *John Stuart Mill and Jeremy Bentham: Utilitarianism and other essays*. New York, NY: Penguin; 2004. 1st published 1861 See chapter 4, 307–14
- Moore, GE. *Principia Ethica*. Amherst, NY: Prometheus Books; 1988. 1st published 1902
- National Human Research Protections Advisory Committee (NHRPAC). [Accessed May 14, 2014] Clarifying specific portion of 45 CFR 46 subpart D that governs children’s research [online]. 2002. <http://www.hhs.gov/ohrp/archive/nhrpac/documents/nhrpac16.pdf>
- Nelson RM. A relative interpretation for minimal risk is unnecessary and potentially harmful to children: Lessons from the Phambili trial. *American Journal of Bioethics*. 2011; 11(6):14–16. DOI: 10.1080/15265161.2011.568588
- Nelson RM, Ross LF. In defense of a single standard of research risk for all children. *Journal of Pediatrics*. 2005; 147:565–66. DOI: 10.1016/j.jpeds.2005.08.051 [PubMed: 16291339]
- Research Involving Children. 43 Federal Register 2083. 1978 Jan 13. Research involving children: Report and recommendations of the national commission for the protection of human subjects of biomedical and behavioral research. Codified at 45 C.F.R. pt. 46
- Resnik DB. Eliminating the daily life risks standard from the definition of minimal risk. *Journal of Medical Ethics*. 2005; 31:35–38. DOI: 10.1136/jme.2004.010470 [PubMed: 15634751]
- Resnik DB. The role of intuition in risk/benefit decision-making in human subjects research. *Accountability in Research*. 2017; 24(1):1–29. DOI: 10.1080/08989621.2016.1198978 [PubMed: 27294429]
- Rid A, Emanuel EJ, Wendler D. Evaluating the risks of clinical research. *JAMA : The Journal of the American Medical Association*. 2010; 304(13):1472–79. DOI: 10.1001/jama.2010.1414 [PubMed: 20924013]
- Rid A, Wendler D. A framework for risk-benefit evaluations in clinical research. *Kennedy Institute of Ethics Journal*. 2011a; 21(2):141–79. DOI: 10.1353/ken.2011.0007 [PubMed: 21696094]
- Rid A, Wendler D. A proposal and prototype for a research risk repository to improve the protection of research participants. *Clinical Trials*. 2011b; 8:705–15. DOI: 10.1177/1740774511414595 [PubMed: 21859783]
- Ross LF. Do healthy children deserve greater protection in medical research? *Journal of Pediatrics*. 2003; 142(2):102–07. DOI: 10.1067/mpd.2003.84 [PubMed: 12584527]
- Rossi J. The prospects for objectivity in risk assessment. *The Journal of Value Inquiry*. 2012; 46(2): 237–53. DOI: 10.1007/s10790-012-9338-9
- Rossi J, Nelson RM. Is there an objective way to compare research risks? *Journal of Medical Ethics*. 2012; 38:423–27. DOI: 10.1136/medethics-2011-100194 [PubMed: 22367000]
- Secretary Advisory Committee for Human Research Protections (SACHRP). [Accessed May 14, 2014] SACHRP chair letter to HHS secretary regarding research recommendations [online]. 2005 Jul 28. <http://www.hhs.gov/ohrp/sachrp/sachrpltrtohhssec.html>
- Shah S. The dangers of using a relative risk standard for minimal risk. *American Journal of Bioethics*. 2011; 11(6):22–23. DOI: 10.1080/15265161.2011.572511

- Shah S, Whittle A, Wilfond B, Gensler G, Wendler D. How do IRBs apply the federal risk and benefit standards for pediatric research? *JAMA : The Journal of the American Medical Association*. 2004; 29:476–82. DOI: 10.1001/jama.291.4.476
- Shrader-Frechette, KS. *Risk and rationality: Philosophical foundations for populist reforms*. Los Angeles, CA: University of California Press; 1991.
- Smith M, Lewis D, Johnston M. Dispositional theories of value. *Proceedings of the Aristotelian Society Supplementary Volumes*. 1989; 63:89–111. 113–137, 139–174. DOI: 10.1093/aristoteliansupp/63.1.89
- Snyder J, Miller CL, Gray G. 'Relative versus absolute standards for everyday risk in adolescent HIV prevention trials: Expanding the debate. *American Journal of Bioethics*. 2011; 11(6):5–13. DOI: 10.1080/15265161.2011.568576
- Stark, L. *Behind closed doors: IRBs and the making of ethical research*. Chicago, IL: University of Chicago Press; 2012.
- Strandberg C. In defence of the open question argument. *The Journal of Ethics*. 2004; 8:179–96. DOI: 10.1023/B:JOET.0000018766.62114.75
- Thompson RA. Vulnerability in research: A developmental perspective on research risk. *Child Development*. 1990; 61:1–16. DOI: 10.2307/1131043 [PubMed: 2307031]
- US Department of Health and Human Services (US DHHS). *Protections for children involved as subjects in research*. 1983 Mar 8. 45 CFR 46 48 FR 9814-20
- US Department of Health and Human Services (US DHHS). *Federal policy for the protection of human subjects*. 1991 Jun 18. 45 CFR 46 FR 56-117
- US Food and Drug Administration (US FDA). [Accessed November 2015] Minutes of Institutional Review Board (IRB) meetings - Draft guidance for institutions and IRBs. <https://www.fda.gov/RegulatoryInformation/Guidances/ucm470046.htm#vote><https://www.fda.gov/RegulatoryInformation/Guidances/ucm470046.htm#vote>
- Wendler D. Protecting subjects who cannot give consent: Toward a better standard for minimal risk. *Hastings Center Report*. 2005; 35(5):37–43. DOI: 10.1353/hcr.2005.0087 [PubMed: 16295263]
- Wendler D. A new justification for pediatric research without the potential for clinical benefit. *American Journal of Bioethics*. 2012; 12(1):23–31. DOI: 10.1080/15265161.2011.634482
- Wendler D, Belsky L, Thompson KM, Emanuel EJ. Quantifying the federal minimal risk standard: Implications for pediatric research without a prospect of direct benefit. *JAMA : The Journal of the American Medical Association*. 2005; 294(7):826–832.s. DOI: 10.1001/jama.294.7.826 [PubMed: 16106008]
- Wendler D, Emanuel EJ. What is a minor increase over minimal risk? *Journal of Pediatrics*. 2005; 147:575–78. DOI: 10.1016/j.jpeds.2005.07.013 [PubMed: 16291344]
- Wendler D, Varma S. Minimal risk in pediatric research. *Journal of Pediatrics*. 2006; 149:855–61. DOI: 10.1016/j.jpeds.2006.08.064 [PubMed: 17137907]
- Westra AE, Wit JM, Sukhai RN, De Beaufort ID. How best to define the concept of minimal risk. *The Journal of Pediatrics*. 2011; 159(3):496–500. DOI: 10.1016/j.jpeds.2011.05.034 [PubMed: 21764402]