Sci Eng Ethics (2008) 14:77–92
DOI 10.1007/s11948-007-9031-8

ORIGINAL PAPER

The Acceptability and the Tolerability of Societal Risks: A Capabilities-based Approach

Colleen Murphy · Paolo Gardoni

Received: 9 January 2007 / Accepted: 24 July 2007 / Published online: 14 September 2007

Abstract In this paper, we present a *Capabilities-based Approach* to the acceptability and the tolerability of risks posed by natural and man-made hazards. We argue that judgments about the acceptability and/or tolerability of such risks should be based on an evaluation of the likely societal impact of potential hazards, defined in terms of the expected changes in the capabilities of individuals. Capabilities refer to the functionings, or valuable doings and beings, individuals are able to achieve given available personal, material, and social resources. The likely impact of a hazard on individuals' capabilities should, we argue, be compared against two separate thresholds. The first threshold specifies the minimum level of capabilities attainment that is *acceptable* in principle for individuals to have in the aftermath of a hazard over any period of time. This threshold captures the level that individuals' capabilities ideally should not fall below. A risk is acceptable if the probability that the attained capabilities will be less than the acceptable level is sufficiently small. In practice, it can be *tolerable* for some individuals to temporarily fall below the acceptable threshold, provided this situation of lower capabilities attainment is temporary, reversible, and the probability that capabilities will fall below a tolerability threshold is sufficiently small. This second, tolerable threshold delimits an absolute minimum level of capabilities attainment below which no individual in a society should ever fall, regardless of whether that level of capabilities attainment is temporary or reversible. In this paper, we describe and justify this Capabilities-based Approach to the acceptability and tolerability of risks. We argue that the proposed theoretical framework avoids the limitations in current approaches to

C. Murphy (✉)
Department of Philosophy, Texas A&M University, 314 Bolton Hall, College Station, TX 77843-4237, USA
e-mail: cmmurphy@philosophy.tamu.edu

P. Gardoni
Zachry Department of Civil Engineering, Texas A&M University, College Station, TX 77843-3136, USA

acceptable risk. The proposed approach focuses the attention of risk analysts directly on what should be our primary concern when judging the acceptability and the tolerability of risks, namely, how risks impact the well-being of individuals in a society. Also, our Capabilities-based Approach offers a transparent, easily communicable way for determining the acceptability and the tolerability of risks.

Keywords Acceptable risk · Tolerable risk · Risk analysis · Capabilities-based approach · Society · Cost-benefit analysis

Introduction

In this paper we argue that a *Capabilities-based Approach* provides a principled foundation for judging the acceptability and tolerability of risks posed by natural hazards (e.g., droughts, earthquakes, fires, floods, hurricanes, tidal waves, tornados, etc.) and man-made hazards (e.g., pollution, terrorist attacks, etc.). Judgments of the acceptability of such risks should be based on an assessment of the likely societal impact of potential hazards, as gauged by the expected impact of a hazard on the capabilities of individuals. Capabilities refer to the functionings, or valuable states of doing or being, individuals are able to achieve given available personal, material, and social resources [1, 2].

In our proposed Capabilities-based Approach, the expected impact of a hazard on individuals' capabilities should be compared against two separate thresholds. The first threshold specifies the minimum level of capabilities attainment that is *acceptable* in principle for individuals to have in the aftermath of a hazard over any period of time. This threshold captures the level that individuals' capabilities ideally should not fall below. A risk is acceptable if the probability that the attained capabilities will be less than the acceptable level is sufficiently small. In practice, it can be *tolerable* for some individuals to temporarily fall below the acceptable threshold, provided this situation of lower capabilities attainment is temporary, reversible, and the probability that capabilities will fall below a tolerability threshold is sufficiently small. The second, tolerability threshold delimits an absolute minimum level of capabilities attainment below which no individual in a society should ever fall, regardless of whether that level of capabilities attainment is temporary or reversible.

There are two important differences between our treatment of the question of acceptable risk and the treatments characteristically found in the literature on acceptable risk. First, in the literature the question of acceptable risk is often answered from the perspective of risks posed by technologies [3, pp. xi, 1–2]. In this paper, we address the acceptability (and tolerability) of risks posed by natural as well as man-made hazards. Second, while it is common to think about the determination of the acceptability of risk as being case-specific [3, p. 3], we argue that there are principled grounds for assessing risks posed by natural and man-made hazards against the same general thresholds.

There are three sections in this paper. The first describes criteria used in the literature on acceptable risk to evaluate particular approaches to acceptable risk,

highlighting their implications for assessing approaches to the acceptability of the risks posed specifically by natural and man-made hazards. The second considers prevailing approaches to acceptable risk and their limitations. The third section introduces a Capabilities-based Approach to acceptable and tolerable risks and shows how the proposed approach avoids the limitations of current approaches.

Criteria for Choosing an Approach to Tolerable and Acceptable Risks

According to a definition widely accepted in the literature on risk analysis, risk refers to a set of *scenarios*, their associated *probability of occurrence* and *impacts* [4]. Very broadly, a risk is acceptable when, based on some criteria or opinions, a society can allow that risk in principle. Different approaches to acceptable risk offer competing understandings of and justifications for various frameworks for determining the acceptability of a risk.

An adequate approach to acceptable and tolerable risk should fulfill the five criteria described below. These criteria are drawn from discussions in the literature on risk by Fischhoff et al. [3, pp. 1–59], May [5], and Murphy and Gardoni [6]. The evaluation of current approaches to acceptable risk and our proposed Capabilities-based Approach should be based on the degree to which such approaches fulfill these criteria.

1. ***An Approach Should Ensure that Relevant Factors are Taken into Account in an Appropriate Way:*** Risk has two constitutive components (*probability* and *impacts*). Judging the acceptability of a risk is fundamentally an evaluative exercise based on certain *value judgments*. Thus, an approach to acceptable risk minimally should include a discussion of probability, impacts, and value judgments [3, pp. 53–55]. In what follows we describe briefly how to account for these components in an appropriate way.
   Assessing Probabilities and Modeling Uncertainties: To properly account for uncertainty, models for risk problems should be probabilistic. Predictions (or forecasting) is required to assess the potential impacts on society of natural and man-made hazards and how likely they are to occur.
   Impacts of Hazard: An approach to the acceptability of risk should evaluate the *expected net societal impacts* of a hazard [5]. The societal impact is relevant because information about a risk and its acceptability is information that policy- and decision-makers may use when deliberating about which mitigation policies to pursue, if any. An approach to acceptable risk should provide guidance in determining which of the many consequences a hazard may have are appropriate to include in gauging the societal impact of a hazard. The inclusions of such consequences should be based on principled considerations.
   Value Judgments: An approach to acceptable risk evaluates the likely societal impact of a hazard. Such evaluations should be made on the basis of criteria which specify, or are based on judgments regarding, values that ought to be protected, promoted, and prioritized. An approach should offer an explanation

of why a particular value or set of values is important and a principled basis for prioritizing competing values in a given way.

2. ***The Required Data Inputs Should be Accurate, Available, and Accessible:*** For the evaluation of risks to be reliable, the evaluation must be based on accurate information concerning the required inputs. The approach must include criteria for measuring and evaluating the quality, completeness, and relevance of inputs. Information can be inaccurate if it is of poor quality, incomplete, or irrelevant.

   An approach to acceptable risk should be practically feasible, given that it addresses a concrete and practical problem, namely, the evaluation of the risks posed by natural and man-made hazards. It should be possible to gather the required inputs; thus, the data inputs must be *available*, given reasonable assumptions of time and resource constraints. The meaning or significance of the data inputs should be understandable and explainable to non-technical experts, and in that sense *accessible*. This ensures that the evaluation of particular risks is feasible [3, pp. 55–56].

3. ***An Approach Should Provide Concrete, Practicable and Theoretically Justified Information and Conclusions on What Types of Action to Take (or not Take):*** Evaluating the acceptability of a specific risk using a particular approach should yield conclusions about whether the risk is acceptable (and so, for example, no mitigation action is required) or unacceptable (and so it is necessary to act to mitigate the potential impacts of a hazard in their magnitude or likelihood of occurrence). Such conclusions should be unbiased (with no systematic error) and objective in the sense that they are derived from a theoretical framework that is justifiable and reliable "(or reproducible) in the sense that repeated application to the same problem should produce the same result" [3, p. 55].

4. ***The Value Judgments and Method of an Approach Should be Transparent:*** Each approach to acceptable risk is formulated on the basis of specific value judgments. These judgments should be *transparent* and easily communicable, so that it is possible for theorists and the general public to critically evaluate them. Transparency of value judgments will also promote general confidence among the public in the decisions formulated from a particular perspective. When the value judgments of an approach are not transparent, both the approach and its specific evaluations are likely to seem questionable and suspect. The purpose being promoted by risk policies becomes clearer and more understandable if the relative weights assigned to particular values are made explicit [5].

   The method for determining the acceptability of risks also should be *transparent*. This makes it possible for theorists and the general public to judge how well a particular framework for acceptable risk has been applied in a specific situation, which helps to ensure that an approach is being used properly. It can also clarify whether the criticisms of a decision about the acceptability of risk are directed at an overall framework or, rather, at the application of a framework in a specific situation.

5. ***An Approach Should Provide a Picture of The Societal Distribution of The Risks:*** To guard against systematic injustice in the distribution of risks, it is important to know who might bear the negative impact and who might benefit in addition to knowing the aggregate risk a society is taking, [5, 7].

An approach should fulfill the above five criteria to generate reliable, informative and theoretically justifiable judgments about the acceptability of risk.

Current Types of Approaches and their Limitations

In this section, we critically discuss four prevailing approaches to acceptable risk. Our overall purpose is to motivate the need for an alternative approach. Current approaches, we show, fail to meet some of the criteria outlined in the previous section. Our discussion in this section sets the stage for our proposed Capabilities-based assessment of acceptable and tolerable risks in the next section.

1. ***Public Judgment:*** According to this approach, a risk is acceptable if it is acceptable to the general public [8]. This view can be seen as a natural requirement of living in a democratic society, where policies are justified in part by expressing the will of the people [8].

   Despite its theoretical appeal, public acceptance is an inadequate guide to the acceptability of risks [5–7, 9, p. 166]. This approach fails to meet the second criteria because the informational basis is not accessible or accurate. Rarely do all sections of a population have access to information about risks and the skills required to interpret that information [8]. Very often individuals have misperceptions about or are indifferent to the risks they face [5, 10]. Public judgments about risk can also be subject to bias [8, 10]. For example, individuals will evaluate the same risk differently, depending on how the risk is presented to them.

   Even if the above problems could be adequately addressed, there remains one fundamental problem with this approach: it does not provide theoretically justified conclusions on what actions to take as required by the third criterion. This approach equates the demands of democracy with some form of majority rule. However, most liberal democratic theorists deny that democracy requires simple majority rule. Individual rights are often used to constrain or shape the democratic process in various ways (for example, constitutional rights are beyond public deliberation and, in fact, frame and guide public deliberation). Democratic politics also involves helping individual citizens develop informed views. Thus, it is a mistake to assume that democracy requires risk analysts to judge risks based on the unformed preferences of such citizens.[1]

[1] This point was made by Douglas MacLean in his keynote address, "Ethics, reasons, and the role of risk analysis," at the conference on "The Ethical Aspects of Risk" hosted by Delft University of Technology June 14–16, 2006.

2. ***Professional or Expert Judgment:*** A natural response to the difficulties with deferring to public judgment is to have technical experts determine what risks are acceptable. Professionals base their judgment in balancing risks and benefits on "professional standards, personal experience, and clients' desires" [3, p. 61]. Professional standards could include those implicitly acquired during technical training and practice as well as others explicitly stated in professional codes.

   There are three main limitations with this approach. First, traditionally professional judgment has tended to be *ad hoc* and has not accounted for different types of uncertainty in a systematic way, thus failing to account for all of the relevant information as required by the first criterion. Second, while experts can more accurately assess actual risks, and in particular the probability of occurrence of specific events, in practice they can be biased in their assessments of these risks. This is especially true when experts base their judgments solely upon their own values or upon their client's desires and ignore or overlook other intuitively relevant factors like the broader good or welfare of society. When relevant factors are ignored, the theoretical justifiability of the practical decisions of experts is undermined and the third criterion for an approach is not met. Third, professionals might not be explicit about what values are motivating their decisions. This failure of transparency, as required by the fourth criterion, can make it difficult to justify those decisions to the public, as required by the fourth criterion, and can contribute to skepticism and cynicism about the decisions made.
3. ***Bootstrapping:*** In the bootstrapping approach, past standards for acceptable risk are used as a guide for present and future decisions about the acceptability of projected risks. This approach assumes that standards will be revised over time in response criticisms or recognized problems, which supports giving past standards normative weight and using them as prescriptive guides [3, p. 100].

   The bootstrapping approach has the virtue of attempting to find standards for risk that are general, not restricted to a particular context or kind of risk. However, there are two main limitations with this approach. First, it lacks a theoretical framework for justifying particular conclusions about acceptable risk as required by the third criterion. Changes, due to technical and scientific progress and social and political transformations, can make one risk acceptable in the past and not in the present or future. For example, with technological progress, for the same scenario the risk may change; there could be lower (or higher) probabilities and smaller (or higher) impacts. In addition, this approach does not take into account the bases upon which previous standards were formulated, which are not always principled considerations. "The bootstrapping methods [...] share the flaw of being subject to the myths, mistakes, and inequities of the society whose decisions they depend on." [3, p. 86].

   The second limitation is the lack of transparency of this approach, as required by the fourth criterion. Following past standards may lead to the right decisions in the present. However, since this approach neither requires nor encourages transparency with respect to how past standards for acceptable risk were formulated, we cannot *show* or *guarantee* this will be the case. This framework lacks principled reasons for thinking that past standards should be a guide to

present standards. We cannot rely on the past being a reliable guide, in part because the origin and rationale behind those standards are not transparent.

4. ***Formal Analysis:*** This is a broadly utilitarian approach [9, p. 164], in which all of the consequences from a particular risk scenario are translated into the same unit of measurement, so that the consequences of taking different courses of action can all be compared. The most common form of formal analysis is cost-benefit analysis [11–13] in which the unit of measure is monetary. This approach adds up the risks and benefits of various courses of action. It measures risks in terms of the amount of money people are willing to pay to avoid, or the compensation they would demand for exposure to, certain risks [13]. Thus, it defines risks and benefits in terms of individuals' subjective preferences, as reflected in market behavior.

   While formal analysis has the virtue of working within or developing a theoretical framework, so that there are principled decisions about acceptable risk, there are three limitations with this approach. The first is that there is no suitable common unit of measurement used. It is difficult to quantify the value of a human life or many of the other costs that might need to be factored into the equation monetarily [11, 5]. Thus, this approach fails to meet the second criterion since the required inputs are in general neither available nor accurate. Second, using the willingness to accept (i.e., pay for) levels of risk or more generally relying on the market to establish the relevant acceptable equilibrium (of cost-benefit tradeoffs) is subject to the same limitations as public judgment [3, p. 94]. Individuals may not be fully knowledgeable of or aware of the relative risks involved when determining what level of income they are willing to trade for exposure to certain risks (failing to meet the first criterion). Third, there is no requirement that the justness of the distribution of risks be considered [9, p. 165, 14], as demanded by the fifth criterion.

Advocates of one of the approaches above may try to further develop each type of approach in response to the objections that have been raised. While we do not deny the possibility of such improvements, our goal in this paper is to introduce a new conceptual paradigm for thinking about acceptable risk, which maintains the strengths of current approaches, avoids their limitations, and has a more solid theoretical justification.

A Capabilities-based Approach and its Benefits

This section describes the initial development of the concept of capabilities in development economics and policy and summarizes our previous application of this framework to risk analysis. We then describe a Capabilities-based Approach to acceptable risk and introduce the notion of tolerable risk. Finally, we show how this approach fulfills the five criteria for an approach to acceptable and tolerable risks.

Background on the Capabilities-based Approach and Application to Risk Analysis

The concept of capabilities was first formulated by Nobel prize-winning economist Amartya Sen and philosopher Martha Nussbaum in the context of development economics and policy, e.g., [2, 15–17]. In development economics and policy, a central question is how to measure the level of development of societies, or the standard of living of individuals. The Capabilities-based Approach defines the standard of living in terms of individual capabilities, or "the ability of people to lead the kind of life they have reason to value." [18]. To understand capabilities, it is first necessary to define *functionings*, which are "valuable acts or [...] states of being," [1, p. 30] that encompass the various things of value an individual does or becomes in his or her life. Examples of functionings include being alive, being healthy, and being sheltered. A *capability* refer to the *positive freedom* of individuals to achieve the corresponding functioning, given his or her available personal, material, social, institutional, and legal resources.

As noted by Sen [2, p. 67], capabilities are distinct from utilities. While capabilities consider the real options available to individuals, utilities are defined in terms of "mental satisfaction" and are measured by considering individual's preferences. So, if an individual chooses A over B, then A has more utility than B. Utilities cannot accurately capture the standard of living of individuals because they fail to acknowledge the phenomenon of adaptive preferences. That is, individuals tend to modify their preferences and expectations based on the circumstances in which they live. So, for example, individuals living in extreme poverty might develop minimal expectations which are easily satisfied. Consequently, from the utilitarian perspective, they can be said to enjoy a high standard of living, despite being objectively deprived. From the Capabilities-based Approach, however, such individuals would not be judged to have a high standard of living because of their lack of fundamental capabilities.

The capabilities framework is used by the United Nations (UN) and development agencies currently to measure the development of societies. Annually, the Human Development Report (HDR) is published by the UN Development Program, which uses tools like the Human Development Index (*HDI*) to measure the level of development of countries. The *HDI* is constructed by identifying three or four capabilities most relevant for development, like the ability of living a long and healthy life, the opportunity for being knowledgeable, and the ability of having a decent standard of living. Since capabilities are abstract concepts and are not directly quantifiable, indicators are then selected to measure in practice the level of individuals' capabilities in a society. Indicators serve as proxies for their corresponding capabilities [19, pp. 140–152]. A Capability Index is then created for each indicator by converting it into a uniform scale. Finally, the *HDI* is formed by combining the Capability Indices. The *HDI* provides "a simple measure [that] is more understandable to the policy-maker and the public, sending a clear message about what makes the measure go up or down" [20].

In previous work [6, 7], we showed how the Capabilities-based Approach provides a theoretical foundation for identifying and quantifying the *net* (positive

and negative) *societal impact* of natural and man-made hazards in risk analysis. In a Capabilities-based Approach, the potential benefits and losses due to a hazard are measured in a consistent way by using *individual capabilities* as the gauge of benefits and losses. Risk refers then to the probability that individuals' capabilities might be reduced. This approach has the benefit of directing the focus of risk analysts on what should be protected, individual capabilities, which are constitutive elements of individual well-being.

In this framework, the overall change in the quality of life of individuals after a hazard is captured by a considering a few properly selected capabilities and their corresponding indicators. A *Hazard Impact Index* (*HII*) can be constructed using the same method as for the *HDI* by comparing the predicted capabilities of individuals after a hazard with the actual capabilities of individuals before.

A Capabilities-based Approach to the Acceptability and the Tolerability of Risks

The intuitive idea behind our proposal for judging the acceptability and the tolerability of risks is the following. The likely societal impact of a hazard, as gauged by its impact on the level of individuals' capabilities, should be compared against two separate thresholds. The first threshold specifies what minimum level of capabilities attainment it is acceptable in principle for individuals to have in the aftermath of a hazard over any period of time. This threshold for acceptability indicates the level of capabilities individuals should ideally not fall below. Some allowance could be made for individuals' capabilities falling below this acceptable threshold during the emergency recovery phase, if the change in capabilities is temporary and reversible. However, there are limits to what is tolerable even during the emergency phase. The second threshold for tolerability captures this limit. It specifies an absolute minimum level of capabilities attainment, below which no individual in a society should ever fall, regardless of whether that level of capabilities attainment is temporary and reversible.

To illustrate these two thresholds, consider the capability of being sheltered, where shelter implies a materially adequate and permanent dwelling. From the Capabilities-based Approach, it is unacceptable for individuals in the aftermath of a hazard to lack permanent and adequate shelter. However, it is tolerable for individuals to have temporary housing arrangements in the emergency phase that immediately follows a hazard. However, it would neither be tolerable nor acceptable for individuals to be left homeless over any period of time or in temporary housing for a prolonged period of time.

With this general description of the distinctions between acceptability and tolerability in mind, we want to discuss in greater detail the process of determining the acceptability and tolerability of a risk. Before we can make any evaluative judgments about the risks natural and man-made hazards pose, we first need information about which capabilities to consider and the expected level at which individuals' capabilities will be in the aftermath of a hazard. This information can

be conveyed through a Hazard Risk Index (*HRI*). We define such an index for a specific hazard i as

$$HRI_i = \sum_{j=1}^{n_C} C_{i,j} \tag{1}$$

where $C_{i,j}$ is the expected value of the indicator for the j-th capability over the considered population and n_C is the number of capabilities considered.

There are five criteria that should be met for a capability to be included in the *HRI*. First, the proposed capability needs to be *important* enough to justify asking others to take actions to advance these capabilities or freedoms. Second, the capability needs to be *influenceable*, i.e., our actions can make a significant difference in whether the capability is realized or not [21]. Third, the choice of capabilities needs to be *directly connected to our overall purpose* in evaluating acceptable risk in the context of natural and man-made hazards. As Sen writes, "The focus has to be on the underlying concerns and values, in terms of which some definable capabilities may be important and others quite trivial and negligible." [1]. For example, if the overarching concern of risk analysts is the safety of individuals, the capabilities selected should be related to how individual safety is affected by hazards. Typically individuals' physical safety and the safety and integrity of their property are impacted by natural and man-made hazards. So, plausible candidates for capabilities could include the capability for living a long life, maintaining bodily integrity, and preserving bodily health. Fourth, in judging the appropriate number of capabilities to consider, N_c, the minimum number of capabilities possible should be chosen (*capabilities parsimoniety*). Finally, each of the capabilities selected should provide information that cannot be ascertained from the other capabilities (*capabilities orthogonality*). These last two conditions ensure that the evaluation of acceptable risk is practically achievable, while at the same time including all the relevant information.

Capabilities themselves are unquantifiable; thus, indicators are needed to measure the level of individuals' capabilities. A first, practical issue is how to choose reliable indicators [7]. In [22], we discuss in detail the kind of empirical information that is required to specify reliable indicators. Here we want to note that the summation for the acceptable risk threshold will be the summation of specific levels of indicators for each capability.

The HRI_i provides an overall picture of the level of capability attainment after a hazard by summing together the level of each individual capability we want to consider, as gauged by indicators. The actual value of the HRI_i after a hazard is unknown before the hazard occurs, so we need to consider the probability of occurrence of each possible value of the HRI_i.

A given hazard might have different degrees of magnitude, M_i, that represent the intensity of the specific manifestation of the hazard. Different magnitudes of a hazard will have different impacts on the capabilities of individuals. We can develop a *Probability Density Function* (PDF), which describes the likelihood of each potential outcome of the HRI_i. For a specific hazard, we can write the

conditional PDF of HRI_i for a certain magnitude of the hazard M_i, as $P(HRI_i \mid M_i)$ where $P(A|B)$ is the conditional probability of having A, given B.

So the probability distribution function of HRI_i can be written integrating out M_i, using the Total Probability Rule [23, pp. 57–59], as

$$P(HRI_i) = \int_{M_i} P(HRI_i|M_i)\, P(M_i)\, dM_i \tag{2}$$

where $P(M_i)$ is the PDF of M_i. This function provides us with information about the likelihood of the impact of hazards of various magnitudes on individual capabilities.

We then compare the predicted values of the HRI_i and their corresponding likelihood against the threshold for acceptable risk. This threshold corresponds to the lowest acceptable value of HRI_i and captures the minimal demands of justice. According to Sen [24], justice requires equal treatment or equal consideration in some important domain, given the moral equality of all individuals. The Capabilities-based Approach defines the domain of equal consideration in terms of capabilities. Justice mandates minimally that all citizens enjoy a certain threshold level of capabilities. As Nussbaum writes, a "necessary condition of justice for a public political arrangement is that it deliver to citizens a certain basic level of capability. If people are systematically falling below the threshold in any of these core areas, this should be seen as a situation both unjust and tragic, in need of urgent attention–even if in other respects things are going well." [16, p. 71]. Claims of justice should be assessed with respect to thresholds for especially important capabilities. So with the capability for being sheltered, this threshold level could mandate some form of permanent (though potentially mobile) housing.

The threshold for acceptable risk will be defined using the same capabilities that constitute the HRI_i. The precise specification of this threshold, which entails describing the acceptable level of attainment for each individual capability, can, as Nussbaum [17] suggests, be the product of internal democratic processes. Nussbaum [17] compares the specification of thresholds to the interpretation of basic rights in a constitutional democracy. Just as court decisions are the way in which rights like the right to free speech get particularized, so too the specification of what precisely chosen capabilities require could be the product of an interpretive process. This interpretive process can be sensitive to the history of each locale and set realistically ambitious aims, given current conditions. To illustrate sensitivity to history, Nussbaum discusses the differences between how the right to free speech is interpreted in the U.S. and in Germany. Whereas the right to free speech does not cover anti-Semitic speech and writing in Germany, it does in the US. Both specifications are reasonable, she argues, and make sense given the history of each locale. To illustrate the process of specifying reasonably ambitious thresholds, Nussbaum discusses India. While it would be unreasonable to constitutionalize a right to college education, given that 35% of the women and 65% of the men currently are literate, it would not be unreasonable to extend to all children the constitutional right to right to primary and secondary education.

The acceptable risk threshold is computed by summing the minimally acceptable levels of the selected capabilities, $C_{acc,j}$, as captured by indicators, and defined as

$$ARI = \sum_{j=1}^{n_C} C_{acc,j} \tag{3}$$

The probability that, for a given hazard i, the HRI_i is bigger than acceptable threshold, $P(ARI \leq HRI_i)$, can be computed using standard reliability methods [25]. Figure 1 shows an illustration of the $P(ARI \leq HRI_i)$ and $P(HRI_i < ARI)$ in relation to the threshold value ARI and the probability distribution of HRI_i. For a complete framework for acceptable risk, it is necessary to determine how small the probability of the HRI_i being below the acceptable threshold must be for a risk to be acceptable. This judgment, like the judgment about the actual threshold for each individual capability, can be determined through internal democratic processes.

In practice, it may not be possible to completely meet the threshold for acceptable risk. It may not be feasible to expect to be able to prevent all citizens from losing their residence or immediately ensuring they have new permanent housing. At the same time, making some allowance for the permissibility of non-acceptable risks should be restricted. Here we want to suggest a way to make principled decisions when faced with the impossibility of maintaining the acceptable level of individuals' capabilities in the aftermath of a hazard.

The key notion we introduce in this context is tolerable risk. The tolerability threshold specifies the level, below the acceptable threshold, that it is tolerable for individual's capabilities to temporarily fall, so long as the lower capability level is temporary and reversible. The urgency of guarding against permanent, unalterable damage to individuals, the environment, or valued structures is much stronger than guarding against easily reversed or remedied impacts. The threshold for the tolerability of risks, TRI, is defined as

$$TRI = \sum_{j=1}^{n_C} C_{\min,j} \tag{4}$$

where $C_{min,j}$ are the minimum levels of the capabilities that should be tolerated at any time and under any circumstance. This second, tolerable threshold outlines an absolute minimum level of capabilities attainment, below which no individual in

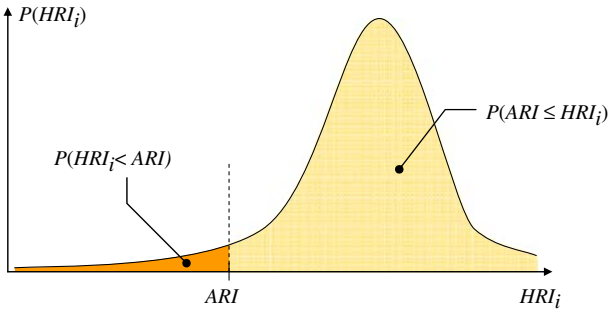

Fig. 1 Illustration of the $P(ARI \leq HRI_i)$ and $P(HRI_i < ARI)$ in relation to the threshold value ARI and the probability distribution of HRI_i

society should *ever* fall. It demarcates what it is to live a minimally human existence. As Nussbaum writes, "there is something to do these functions in a truly human way, not a merely animal way." [16, p. 72]. With the example from shelter, there would be a form of shelter that is appropriate for human beings. The probability should be sufficiently small that a human being will ever fall below this tolerable threshold, regardless of her situation.

In summary we can write that

$$\begin{aligned}
&\text{Probability of being acceptable} && P(ARI \le HRI_i)\\
&\text{Probability of not being acceptable but being tolerable}\\
&\quad\text{(Under condition of reversibility and temporariness)} && P(TRI \le HRI_i < ARI)\\
&\text{Probability of not being tolerable} && P(HRI_i < TRI)
\end{aligned} \tag{5}$$

A similar formulation and discussion could also be followed using the probabilities of exceeding acceptable and tolerable thresholds for each individual capability, rather than simply looking at the aggregated thresholds *ARI* and *TRI*. One concern with simply summing up the capabilities is that a high level of attainment of one capability, above the acceptable threshold or tolerable threshold, might compensate a level of attainment below the tolerable or acceptable threshold of another capability. However, since each capability is incomparable, no amount of reduction in one capability can offset a corresponding gain in another capability. So, it is reasonable to want to ensure that the level of all of the capabilities we consider is above the corresponding threshold mark. In this case the impacts of a hazard would be (a) acceptable if each capability is above the corresponding acceptable level, (b) not acceptable but tolerable when at least one capability is below the corresponding acceptable level but they are all above the minimum tolerable level, and (c) not tolerable if at least one capability is below the minimum tolerable level. So we can write that

$$\begin{aligned}
&\text{Probability of being acceptable} && P\left[\bigcap_j \left(C_{acc,j} \le C_{i,j}\right)\right]\\
&\text{Probability of not being acceptable}\\
&\quad\text{but being tolerable} && P\left\{\left[\bigcap_j \left(C_{\min,j} \le C_{i,j}\right)\right] \cap \left[\bigcup_j \left(C_{i,j} < C_{acc,j}\right)\right]\right\}\\
&\text{Probability of not being tolerable} && P\left[\bigcup_j \left(C_{i,j} < C_{\min,j}\right)\right]
\end{aligned} \tag{6}$$

Finally, our discussion to this point has not explicitly considered the distribution of risks. One of the criteria for an adequate approach is that it provides information about the distributions of risks. Underlying the importance of the distributions is a concern of justice. Frequently, some groups in a society are systematically subject to higher risks than others. It intuitively seems unjust if only some individuals in a community are at risk, especially when those individuals are not sharing in any of

the benefits that come from the risks. To determine whether such systematic disparities exist within a society, we can use a method of disaggregation. Such a method allows us to determine and evaluate the distribution of risks among socio-economic, racial, geographic, and occupational groups. A similar method is used by the United Nations to assess the level of development of sub-groups within a society, considering, for example, geographical, ethnic, and gender groups [26, p. 136]. We can disaggregate the society by groups or regions and see whether there are differences in the risks associated with a hazard. We want the level of protection to be the same for all people. While the total HRI_i within a society is above the *ARI*, one needs to guarantee that no sub-group is systematically below the *ARI*.

Benefits of the Capabilities-based Approach

The proposed Capabilities-based Approach fulfills the criteria for a framework for acceptable risk outlined above. First it accounts for all relevant factors, including the probabilistic nature of the problem. Like the formal analysis, the selection of the relevant impacts to consider is made based on principled considerations. However, the information basis for the Capabilities-based Approach is different. Rather than defining the costs and benefits in terms of preferences or monetary value, this approach focuses on the capabilities of individuals. Thus, it provides a formulation for capturing the actual net societal impacts of a hazard in a way that focuses our attention on what we most care about, namely, how the lives of individuals are affected by a hazard. Similarly, it directs the attention of risk analysts directly on what should be our primary concern when it comes to the acceptability and tolerability of risks, namely, how risks impact the well-being of individuals in a society.

Second, the required data inputs are accurate, available and accessible. Indicators, selected according to criteria outlined in [7], allow us to measure accurately the level of capabilities of particular individuals. As the work in human development has shown, the information from indicators can be internationally available. The data inputs from the framework are also accessible because capabilities capture levels of well-being along dimensions with which all individuals have some familiarity.

Third, the Capabilities-based Approach provides concrete, practicable and theoretically justified information and conclusions on what types of action to take (or not take). The thresholds for acceptability and tolerability provide determinate information regarding whether any action should be taken, based on a theoretically justified framework.

Fourth, the underlying value judgments and method of this approach are explicit and easily communicable to the public. The overall value commitment this approach uses as its starting point is the idea of improving and protecting the well-being of individuals globally. This transparent and intuitively compelling value judgment will inspire confidence in the decisions made from this framework. It also can facilitate critical debate. The method used to analyze the acceptability and tolerability of risk is also transparent. Such judgments are made on the basis of

public, general standards for acceptable or tolerable capability attainment. This transparent analysis of risk gives people the ability to assess the method and its resulting judgments and to argue against them, if they believe they are wrong. In this sense, the transparency of the Capabilities-based Approach might be truer to democratic values than crude measures of public judgment.

Fifth, we outlined a method for providing a picture of the societal distribution of the risks, based on the disaggregation of the HRI_i. Such approach permits us to assess if there is an unequal distribution of risks, for example, among specific geographical, ethnic, and gender groups.

Conclusions

In this paper, we propose a new framework for thinking about the acceptability and the tolerability of risk. Societal risks from natural and man-made hazards are defined in terms of individual capabilities. We define a Hazard Risk Index (*HRI*) that captures the overall impact of a hazard on individual capabilities. A risk is acceptable if individual capabilities remain above a specified threshold. Risks that are voluntary and with reversible impacts are defined as tolerable so long as individual capabilities remain above an absolute minimum threshold. An argument for the selection of the acceptable and tolerable levels of risk is provided and a proper formulation is presented to account for the uncertainty inherent in the risk analysis. Five criteria for a correct formulation of acceptable risk have been presented. The proposed method satisfies each of the proposed criteria.

Acknowledgments A draft of this paper was presented at the conference on the "Ethical Aspects of Risk" at Delft University of Technology in Delft, The Netherlands, June 14–16, 2006. The authors are grateful for the very helpful comments they received, in particular from Profs. Douglas McLean (Department of Philosophy, University of North Carolina at Chapel Hill) and Paul Slovic (Department of Psychology, University of Oregon). We also thank Ms. Katya Hosking and Mrs. Kathleen Murphy for their helpful suggestions.

References

1. Sen, A. (1993). Capability and well-being. In M. Nussbaum & A. Sen (Eds.), *The Quality of Life* (pp. 30–53). Oxford, United Kingdom: Clarendon Press.
2. Sen, A. (1999). *Development as freedom*. New York: Anchor Books.
3. Fischhoff, B., Lichtenstein, S., Slovic, P., Derby, S. L., & Keeney, R. L. (1981). *Acceptable risk*. Cambridge, United Kingdom: Cambridge University Press.
4. Kaplan, S., & Gerrick, B. J. (1981). On the quantitative definition of risk. *Risk Analysis, 1*, 11–27.
5. May, P. J. (2001). Societal Perspectives about earthquake performance: The fallacy of "Acceptable Risk". *Earthquake Spectra, 17*(4), 725–737.
6. Murphy, C., & Gardoni, P. (2006). The role of society in engineering risk analysis: A capabilities-based approach. *Risk Analysis, 26*(4), 1085–1095.
7. Gardoni, P., & Murphy, C. (2006). A capabilities-based approach to measuring the societal impacts of natural and man-made hazards. *Natural Hazard Review* (Submitted).
8. Hunter, P. R., & Fewtrell, L. (2001). Acceptable risk. In L. Fewtrell & J. Bartram (Eds.), *Water quality: Guidelines, standards, and health* (pp 207–227). London, UK: IWA Publishing.

9. Harris, C. E., Pritchard, M. S., & Rabins, M. J. (2005) *Engineering ethics: Concepts & cases*, 3rd ed. Belmont: Wadsworth Press.
10. Slovic, P., Fischhoff, B., & Lichtenstein, S. (1979). Rating the risks. *Environment, 21*(3), 36–39, 14–20.
11. Fischhoff, B., Slovic, P., & Lichtenstein, S. (1979). Weighing the risks: Which risks are acceptable. *Environment, 2*(4), 17–20, 32–38.
12. MacLean, D. (ed.) (1986). *Values at risk*. Towata, NJ: Rowman and Allanheld.
13. Anderson, E. (1988). Values, risks, and market norms. *Philosophy and Public Affairs, 17*(1), 54–65.
14. Paté, M. E. (1983). Acceptable decision processes and acceptable risks in public sector regulations. *IEEE Transactions on Systems, Man, and Cybernetics, 13*(3), 113–124.
15. Sen, A. (1989). Development as capabilities expansion. *Journal of Development Planning, 19*, 41–58.
16. Nussbaum, M. (2000a). *Women and human development: The capabilities approach*. New York: Cambridge University Press.
17. Nussbaum, M. (2000b). Aristotle, politics, and human capabilities: A response to Antony, Arneson, Charlesworth, and Mulgan. *Ethics, 111*(1), 102–140.
18. Anand, S., & Sen, A. (2000). The income component of the human development index. *Journal of Human Development, 1*(1), 83–106.
19. Raworth, K., & Stewart, D. (2003). Critiques of the human development index: A review. In S. Fukuda-Parr & A. K. Shiva Kumar (Eds.), *Readings in Human Development* (pp. 140–152). Oxford, UK: Oxford University Press.
20. Fukuda-Parr (2003). The human development paradigm: Operationalizing Sen's ideas on capabilities. *Feminist Economics, 9*(2–3), 301–317.
21. Sen, A. (2004). Elements of a theory of human rights. *Philosophy & Public Affairs, 32*(4), 315–356.
22. Gardoni, P., Murphy, C., & Sanchez, M. (2006). The Implementation of a capabilities-based approach to measuring the societal impacts of natural and man-made hazards. *Risk Analysis* (Submitted).
23. Ang, A. H-S., & Tang, W. H. (2006). *Probability concepts in engineering*. Hoboken, NJ: John Wiley & Sons.
24. Sen, A. (1996). On the status of equality. *Political Theory, 24*(3), 394–400.
25. Ditlevsen, O., & Madsen, H. O. (1996). *Structural reliability methods*. New York: Wiley.
26. Jahan, S. (2003). Evolution of the human development index. In S. Fukuda-Parr & A. K. Shiva Kumar (Eds.), *Readings in human development* (pp. 128–139). Oxford, UK: Oxford University Press.