December 20 1996

Insensitivity to the Value of Human Life: A Study of Psychophysical Numbing

David Fetherstonhaugh*
Department of Psychology
Stanford University
Stanford, California 94305-2130
(415) 725-5487; fax (415) 725-5699;
email df@psych.stanford.edu

Paul Slovic
University of Oregon and
Decision Research, Eugene, Oregon

Stephen M. Johnson
Decision Research

James Friedrich
Willamette University

Journal of Risk and Uncertainty, in press

Abstract

A fundamental principle of psychophysics is that people's ability to discriminate change in a physical stimulus diminishes as the magnitude of the stimulus increases. We find that people also exhibit diminished sensitivity in valuing lifesaving interventions against a background of increasing numbers of lives at risk. We call this "psychophysical numbing." Studies 1 and 2 found that an intervention saving a fixed number of lives was judged significantly more beneficial when fewer lives were at risk overall. Study 3 found that respondents wanted the minimum number of lives a medical treatment would have to save to merit a fixed amount of funding to be much greater for a disease with a larger number of potential victims than for a disease with a smaller number. The need to better understand the dynamics of psychophysical numbing and to determine its effects on decision making is discussed.

Key Words: Decision making, life saving, value of life, risk-benefit analysis, psychophysical numbing

Nobelist Albert Szent-Gyorgi once observed, "I am deeply moved if I see one man suffering and would risk my life for him. Then I talk impersonally about the possible pulverization of our big cities, with a hundred million dead. I am unable to multiply one man's suffering by a hundred million." Most people seem to at least tacitly appreciate the kind of insensitivity toward loss of human life articulated in Szent-Gyorgi's statement. We recognize the need for creative attempts to drive home the severity of catastrophic losses. One activist group lobbied Congress by placing 38,000 pairs of shoes, boots, and sneakers around the Capitol building to sensitize representatives to the 38,000 gunshot fatalities America experiences annually ("38,000 Shoes," 1994). Another example is given by Rummel (1995), who asked people to consider this century's total democide (state sanctioned killing, aside from warfare) of 170,000,000 by imagining a chain of bodies laid head to toe reaching from Honolulu, across the Pacific and the continental U.S., to Washington D.C. and then back again more than 16 times. Losses of life framed in these ways attempt to mitigate the insensitivity that seems to occur so naturally when we try to comprehend past tragedies or think rationally about how to mitigate or prevent large losses of life in the future.

What psychological principles lie behind this insensitivity? In the 19th century, E. H. Weber and Gustav Fechner discovered a fundamental psychophysical principle that describes how we perceive and discriminate changes in our physical environment. They found that people's ability to detect changes in a physical stimulus decreases rapidly as the magnitude of the stimulus increases (Weber, 1834; Fechner, 1860). What is known today as "Weber's law" states that in order for a change in a stimulus to become just noticeable, a fixed percentage must be added. Thus perceived difference is a relative matter. To a small stimulus, only a small amount must be added. To a large stimulus, a large amount must be added to be equally noticeable. Fechner proposed a logarithmic law to model this nonlinear growth of sensation. Numerous empirical studies by S. S. Stevens (1975) have demonstrated that the growth of sensory magnitude (ψ) is best fit by a power function of the stimulus magnitude ϕ

$$\psi = k\phi^{\beta}$$

where the exponent β is typically less than one for measurements of phenomena such as loudness, brightness, and even the value of money (Galanter, 1962).[1] For example, if the exponent is 0.5 as it is in some studies of perceived brightness, a light that is four times the intensity of another light will be judged only twice as bright.

Our cognitive and perceptual systems seem to be designed to sensitize us to small changes in our environment, possibly at the expense of making us less able to detect and respond to large changes. As the psychophysical research indicates, constant increases in the magnitude of a stimulus typically evoke less and less of a change in response. Applying this principle to the valuing of human life suggests that a form of psychophysical numbing may result from our inability to appreciate losses of life as they become more catastrophic—a phenomenon that could impair our ability to make consistent, equitable, and wise decisions.[2]

Evidence of psychophysical numbing comes from a study by Summers, Slovic, Hine, and Zuliani (in press), who hypothesized that people may exhibit a systematic distortion in perception of death tolls from wars not unlike the systematic distortion found in many traditional experiments in sensory psychophysics. They found that deaths from wars were perceived according to a power function where $\beta = 0.32$. Thus, respondents in these experiments perceived a war that claimed nine times the number of lives as a second war to be only about three times greater in magnitude. The degree of psychophysical numbing changed in these experiments as a function of how the losses were framed. Respondents' insensitivity was reduced ($\beta = .99$) when the same total number of casualties was presented as "deaths per day" rather than "deaths per war."

Kahneman and Tversky (1979) have incorporated this psychophysical principle of decreasing sensitivity into prospect theory, a descriptive theory of decision making under uncertainty. A major element of prospect theory is the value function, which relates subjective value to actual gains or losses. The function is concave for gains and convex for losses. When applied to human lives, the value function implies that the subjective value of saving a specified number of lives is greater for a smaller tragedy than for a larger one (when the life-saving effort is framed as reducing a loss). Such psychophysical numbing may have dramatic implications for the judgments and decisions people make. For example, an intervention that reduces the number of deaths in a tragedy from 2000 to 1000 may be judged substantially more valuable than one that reduces deaths from 99,000 to 98,000. Even though both interventions save the <u>same</u> number of lives, in the former people may decide to act while in the latter they may not, perhaps under the impression that saving 1,000 lives out of 2,000 is a significant proportion but saving 1,000 out of 99,000 is merely "a drop in the bucket."

How should we value the saving of a life? We believe that, in most circumstances, "a life is a life"—the value of saving a certain number of people from death should not be affected by the number or proportion of others who remain unsaved. This perspective presumes a linear relationship between the number of lives one can save in a given situation and the value associated with saving them. Thus an effort saving 200 lives would have twice the value of another that saves 100 lives in the same circumstances. This would lead decision makers to prefer the intervention that saves the greatest number of lives even if that number is proportionally smallest when compared to the number at risk. Stated differently, we argue that the value of lives saved should be based on the number an intervention can save, and should therefore be independent of the size of the population from which the saved lives originate.

Under a one-to-one correspondence between the number and value of saved lives, the value of a life-saving effort should also be independent of when in the process those lives are saved. For example, the value of saving the first 100 individuals in a tragedy should not change if instead these individuals happened to be the last 100 saved. According to prospect theory's curved value function, however, the value of saving lives will in many cases depend on when in the process those lives are saved. For example, the value of reducing deaths by 100 early in an intervention would not likely be equivalent to that of an identical reduction later in the process; a reduction in loss of life that brings the death toll closer to zero might appear more valuable.

Except for the study by Summers, Slovic, Hine, and Zuliani (in press) and a demonstration by Tversky and Kahneman (1981) showing that the nonlinear value function does seem to apply to gains and losses of life, little empirical work has been conducted that investigates psychophysical numbing in the domain of life saving. The three studies reported here explored how people value life-saving interventions. We hypothesized that respondents' judgments would exhibit psychophysical numbing by responding to life-saving interventions in a manner consistent with prospect theory's value function. Studies 1 and 2 examined how the perceived benefit of saving lives changed when interventions saving the same number of people are implemented in tragedies that differ in magnitude. We predicted that such life-saving interventions would be valued more highly when the number of lives at risk was small than when the number at risk was large. We also predicted that saving lives later in an intervention, bringing the death toll closer to zero, would be valued more highly. Study 3 examined how the total number of people at risk influenced people's estimates of the number of lives an intervention must save to justify a fixed amount of funding. We predicted that, when the number at risk was larger, the intervention would be required to save more lives.

1. Study 1

1.1. Method

Materials and Procedure. Undergraduate volunteers (n = 54) from two sections of an economics statistics course were instructed in a short questionnaire to imagine themselves as a government official of a small, developing country and were asked to evaluate four government programs (Programs A, B, C, and D) being considered for funding. Each of the programs "cost about the same" and addressed the following issues: the employment problem in their country, the transportation problem in their country, and the life-threatening refugee problem in Rwanda.[3] The transportation program proposed to remedy poor road conditions, and the employment program proposed to decrease the jobless rate. There were two Rwandan refugee programs, each proposing to provide enough clean water to save the lives of 4,500 refugees suffering from cholera in neighboring Zaire. The Rwandan programs differed only in the size of the refugee camps where the water would be distributed; one program proposed to offer water to a camp of 250,000 refugees and the other proposed to offer it to a camp of 11,000.

Respondents evaluated the programs in pairs, one pair per page. Because the two Rwandan programs were never paired together, only five of the six possible pairings appeared in the booklets. All respondents evaluated the same paired comparisons, presented in one of two randomized orders. Each page contained brief descriptions of two programs being compared, followed by a response scale such as that shown below for Programs A and B.

Program A														Program B
	6	5	4	3	2	1	0	1	2	3	4	5	6	
	Strong preference for A				Slight preference for A		No pref.	Slight preference for B					Strong preference for B	

On the last page, participants responded to several questions designed to verify whether they perceived that the same number of refugees would be saved by either of the Rwandan programs. The final item requested respondents to briefly explain whether it was better to save lives in the smaller or the larger refugee camp, and why.

1.2. Results and Discussion

The manipulation checks verified that most respondents correctly perceived that the two Rwandan programs saved the same number of lives.

Ratings on the 13-point preference scale constituted the dependent measure. For the four pairings containing a Rwandan program, participants' responses were subsequently recoded so that a preference for the Rwanda program in a pair was indicated by a positive number and a preference for the non-Rwandan program in a pair was coded as a negative number. For example, in a pair containing the large-camp program and the transportation program, if a participant circled a "2" to indicate a slight preference for the transportation program, the rating would have been re-coded as a "–2." Thus, participants' re-coded responses ranged from –6 to +6. Because responses in the transportation vs. employment program comparison were not of theoretical interest, they were excluded from the analysis. An analysis of variance on respondents' preferences revealed no effects due to respondents' gender or to the order in which the paired comparisons were presented. Data were therefore combined without regard for these variables.

We predicted that preference ratings would be greater for the small-camp program than the large-camp program. Because these programs were never paired together, however, we compared respondents' ratings for the two Rwandan programs in pairings that shared a common non-Rwandan program. For example, we compared respondents' ratings in the transportation vs. the small-camp pairing with their ratings in the transportation vs. the large-camp pairing. We expected that the re-coded rating for the small-camp pairing would be greater than the rating for the large-camp pairing.

This prediction was tested using a within-subjects, 2 x 2 analysis of variance on preference ratings, varying comparison program type (transportation or employment) and Rwanda camp size (large or small). As predicted, the results revealed a camp-size main effect, $F(1, 52) = 8.24$, $p < .01$ (see Figure 1). Even though most respondents realized that the same number of refugees could be saved in either camp, they preferred the small-camp program ($M = .45$) over the large-camp program ($M = -.20$) when paired with either the transportation or employment programs.

The same ANOVA was conducted on the preferences of 22 respondents who indicated on the last page of the booklet that saving 4,500 lives in the large camp was neither better nor worse than saving 4,500 lives in the small camp. Even these respondents, who indicated no preference between one life-saving Rwandan program and the other when asked directly, preferred the small-camp program ($M = .93$) over the large-camp program ($M = .41$) when evaluations were masked by paired comparisons, $F(1, 21) = 3.92$; $p = .06$.

One last question asked respondents whether it was better to save lives in the smaller or the larger refugee camp, and to state why. About 44% of respondents reported that it was better to save lives in the smaller camp. As perhaps the strongest evidence for the psychophysical numbing hypothesis thus far, this result is quite remarkable, especially considering that the life-saving potential of each Rwandan program was reinforced by the preceding question in which nearly all respondents reported that the interventions would save the same number of lives. Approximately 42% of respondents reported no preference between the two programs and 14% indicated that it was better to save lives in the larger camp.

We suggest that the sizable proportion of respondents who preferred to save lives in the smaller camp reflects people's general tendency to become desensitized to the life-saving potential of interventions applied to larger tragedies. However, what appears to be psychophysical numbing might not be caused by insensitivity at all, but actually by respondents' sensitivity to preventing further casualties among refugees—an eminently reasonable goal. Respondents might have preferred the small-camp program because of the increased hazard of administering a limited supply of a scarce commodity to a large and desperately needy group of people, as might likely be found in a large camp. Not only might such an effort incur additional casualties through the riot it could spark, but people could be at greater risk of infection or later re-infection due to the increased tendency for water-borne diseases to spread in a larger compared to a smaller refugee camp. We have labeled these explanations the riot and contagion hypotheses, respectively.

Evidence in the present study suggested that most respondents considered neither the riot nor the contagion hypotheses. In fact, only one participant mentioned the riot hypothesis as a rationale when responding to the final question (i.e., whether it was better to save lives in the small or the large refugee camp, and why). Seven respondents, however, did cite some form of the contagion hypothesis, though several used it to support their preference for saving lives in the larger camp. The fact that even the 22 respondents who stated that saving lives in the larger camp was neither better nor worse than saving lives in the small camp exhibited psychophysical numbing also speaks against these hypotheses. Unless a substantial portion of respondents used but failed to report one or both of these hypotheses as part of their rationale, we believe it unlikely that either hypothesis could be widely responsible for the effects found in the above analyses.

Furthermore, the next study reports data that not only replicates the present study, but provides evidence that essentially rules out both the riot and the contagion hypotheses as alternative explanations for psychophysical numbing.

2. Study 2

Study 2 retained much of the content and structure of the previous study. Participants first read a cover story about the Rwandan refugee crisis and then evaluated one small country's life-saving intervention proposed for several refugee camps. For each camp, all respondents answered two questions: (1) how beneficial would sending the aid be? and (2) should aid be sent or not? Study 2, however, differed from the previous study in several respects. In Study 2, comparisons between the Rwandan scenarios were easier, which would presumably lessen psychophysical numbing among respondents. Whereas Study 1 paired each Rwandan scenario with a "dummy" scenario, making direct comparisons between Rwandan scenarios more difficult, Study 2 omitted dummy scenarios and had respondents evaluate Rwandan scenarios individually.

Though easier in this regard, Study 2 was generally more complex because it contained a more detailed cover story and incorporated two additional independent variables. Besides the camp-size manipulation found in the previous study, Study 2 manipulated when in the life-saving process the humanitarian aid was distributed. We predicted that saving a portion of lives near the end of a crisis would be valued more highly than saving an equal portion near the beginning of a crisis because the former solves virtually all the problem whereas the latter does not. Study 2 also manipulated the "reliability" of the equipment used to administer the aid (i.e., purified water). We included this variable to discourage respondents from rating the intervention as maximally beneficial in every scenario.

2.1. Method

Overview of Design. The present study manipulated three within-subjects variables: size of refugee camp (11,000 or 250,000), amount of pure-water aid a camp was receiving before a water-purification plane was sent (low or high), and reliability of the plane (60% or 100%). This yielded the eight different scenarios participants read and it allowed us to analyze their responses in a $2 \times 2 \times 2$ repeated-measures factorial design.

All respondents evaluated the same eight scenarios. Half received the block of four 100%-reliable plane scenarios first and the block of four 60%-reliable plane scenarios second, and half received the blocks of four in the reverse order. Within each block of four scenarios, the "camp-size" and "prior-help" variables were mixed according to a latin-square design.

There were two dependent variables: (1) the rated benefit of sending a plane, and (2) a yes/no decision on whether or not to send a plane.

Materials and Procedure. University of Oregon students (n = 162) were paid $4 to complete an 11-page questionnaire about the Rwandan refugee crisis. The cover story of the questionnaire informed respondents that the U.N. High Commissioner for Refugees was coordinating a massive humanitarian aid campaign by requesting that able countries send assistance to the Rwandan refugees in Zaire. Many refugees had a water-borne disease and would die if purified water did not soon become available. One small country was considering sending one of two Dash-8 water-purification planes to Zaire. Although each water system was capable of producing only a small fraction of the water needed, each could keep about 1500 disease victims alive each day. The purification system in one plane was 100% reliable, and the system in the other plane was only 60% reliable—reliable in the sense that there was only a "60% chance that the system would work once it got to Zaire." Once a plane was operating in a camp, respondents were informed that, "aid-workers will distribute the clean water to designated disease victims, which usually saves the victims' lives." The cost to this small country of delivering and operating these purification systems was significant in light of its economy.

The following pages contained eight scenarios about the four refugee camps (see Table 1 for a summary of information given in the eight scenarios). Each scenario was identically structured. For example, on one page respondents read the following scenario (Scenario 1):

> "The city of Moga in Zaire now has about 11,000 Rwandan refugees. Few water purification systems from other countries are now in place. 5% of the clean water needed for disease victims in this camp is currently being met. If the 100%-reliable Dash-8 water

purification plane is sent to Moga, 50% of this camp's water need for disease victims would be met."

For scenarios using the 60%-reliable plane, the following phrase was added: ". . . , provided the purification system works."

In the other small-camp scenarios (Fizi 1 and Fizi 2—see Table 1), 50% of the clean water need was currently being met, so the aid increased this to 95%, provided the system worked.

In the two scenarios involving the large (250,000 refugees) camp, Uvira, the prior aid met 5% of the water need and the additional aid would bring this to 7%. In the scenarios involving the Kalehe camp (scenarios 7 and 8), 93% of the water need was being met and the additional aid would bring this to 95%. Thus, the intervention proposed to save 2% of disease victims in a given large camp and 45% of disease victims in a given small camp. Recall, however, that the same absolute number of lives (1500) would be saved in each case, regardless of camp size.

Each scenario was followed by two questions. First, "What would be the benefit of sending this Dash-8 plane to this camp?" Respondents answered this question on a nine-point Likert scale, titled "Benefit," anchored at the ends by: 0 ("extremely low benefit") and 8 ("extremely high benefit"). Second, they were asked, "Given the benefit indicated on the scale above, would it be worth sending the plane to this camp?" Respondents circled either "Yes" or "No." On each page, participants were reminded that responses to each scenario should be independent of their responses to the other scenarios.

After completing this task, participants responded to a question designed to assess whether they correctly perceived that the same number of lives would be saved by an intervention, regardless of the size of camp where it was implemented.

2.2. Results and Discussion

A check on subjects' understanding of the problem revealed that 60% of respondents correctly perceived that the water systems would save about the same number of lives regardless of refugee camp size, 23% believed that substantially more lives would be saved in the larger camp, and 17% believed that substantially more lives would be saved in the smaller camp. The analyses reported below omitted this last group of respondents because their belief could have quite reasonably lead them to prefer implementing the intervention in the small camps, not because of psychophysical numbing but simply because it could have saved more lives.

A 2 × 2 × 2 within-subjects ANOVA on respondents' benefit ratings provided strong support for the psychophysical numbing hypothesis (see Figure 2). A significant main effect for camp size, $F\ (1, 132) = 160.5$, $p < .001$, indicated that respondents believed sending the planes to small camps was more beneficial ($M = 6.46$) than sending them to large camps ($M = 4.54$). A main effect for the prior-aid variable, $F\ (1,132) = 15.35$, $p < .001$, indicated that respondents believed sending the planes to camps that were already satisfying a substantial portion of their clean-water need was more beneficial ($M = 5.73$) than sending them to camps that were only satisfying a small portion of their water need ($M = 5.27$). And, not surprisingly, the results revealed a main effect for plane reliability, $F\ (1, 132) = 12.01$, $p < .001$, indicating that respondents believed the 100%-reliable plane ($M = 5.67$) was more beneficial than the 60%-reliable plane ($M = 5.33$). No other effects were significant.

As predicted, respondents appeared to favor interventions more when implemented in the later stages of the life-saving process. For example, respondents thought it was more beneficial to save 2% of those at risk when the threat of a tragedy was nearly contained than when it was just beginning to take its toll. As the absence of interaction effects indicated, this was as true for small camps as it was for large.

A 2 × 2 × 2 ANOVA on respondents' dichotomous decisions about whether or not to send the planes to the camps also revealed a significant main effect for camp size, F (1, 130) = 105.4, $p < .001$, indicating that respondents decided to send the planes to small camps more often (93%) than to large camps (59%). A main effect for plane reliability, F (1, 130) = 4.61, $p < .05$, indicated that respondents decided to send the reliable plane to the camps slightly more often (78%) than the unreliable plane (74%). Interestingly, the main effect for prior aid was not significant, F (1, 130) = .47, $p = .50$. Respondents decided to send the planes 75% of the time to camps receiving little prior aid, and 77% of the time to camps receiving substantial prior aid. No other effects were significant.

The above analyses show that respondents' judgments and decisions about sending aid to refugee camps differed greatly depending on camp size. It is possible that such responses could be justified if they were based on some rationale such as the riot hypothesis or the contagion hypothesis. However, in addition to the evidence against the riot and contagion hypotheses from Study 1, data from the present study provided strong evidence that neither of these alternative hypotheses were widely considered. If respondents had considered such explanations, one would have expected them to substantially devalue the interventions for those scenarios in which the risk of rioting or re-infection was the greatest, namely, large camps in the early stages of the life-saving process. In these scenarios, respondents were faced with implementing an intervention whose supply of aid was particularly inadequate for the demand (see Table 1, Uvira 1 and Uvira 2). The data, however, did not reflect this interaction pattern. Rather, respondents devalued interventions to the same degree both in large camps needing little additional aid and large camps needing massive additional aid.

3. Study 3

In the previous studies, we asked respondents to make evaluations about one type of intervention (saving a fixed number of lives) applied to several tragic circumstances varying in magnitude. In Study 3 we asked respondents to estimate the minimum number of lives each of several interventions must save to merit a fixed amount of money. If people tend to view an amount of assistance in a large tragedy as less valuable than an equivalent amount in a small tragedy, as was shown in the previous studies, then they should require more life-saving assistance to be "added" to the large tragedy to make the assistance in each of equal value. In the present study, therefore, we predicted that respondents' estimates of the minimum number of lives each intervention would have to save would be greater for the larger than the smaller tragedies.

3.1. Method

Materials and Procedure. University of Oregon students ($\underline{n} = 165$) were paid $4 to complete a questionnaire asking them to imagine that they were the chairperson on the board of "Science For Life," a charitable foundation in charge of distributing large sums of money to research institutions that develop treatments for serious diseases. Each respondent was asked to determine which medical institutions Science For Life should fund with its limited resources.

The three medical institutions (X, Y, and Z) that were requesting support each proposed to implement a new treatment that would significantly reduce the annual number of deaths caused by a particular disease.

Respondents were also instructed to assume that: (1) the treatments will induce a cure for some people and thus "save their lives," and (2) the people who are not cured will experience no other beneficial effect; that is, the treatment will not improve their "quality of life."

Respondents completed two tasks: an estimation task and a ranking task. The first task required them to estimate for each disease "How large a reduction in yearly deaths makes [the] institution worthy of funding?"

Each respondent made estimates for all three medical institutions. Each page of the questionnaire presented information about one medical institution, X, Y, or Z, and information about the number of deaths caused in the previous year by the disease for which the institution proposed treatment, Disease A, B, or C, respectively (see "Task 1 Information" in Table 2). Thus, on each page respondents read the following:

"Medical Institution (X) [Y] {Z} has developed a treatment for Disease (A) [B] {C} and now requests $10 million from Science For Life. Last year, people with Disease (A) [B] {C} did not have access to this treatment, and (15,000) [160,000] {290,000} died from the disease. Given Science For Life's shrinking budget, what is the minimum number of lives this treatment would have to save next year in order for Medical Institution (X) [Y] {Z} to merit funding?"

Respondents recorded their estimates on a blank line provided on each page. Six versions of the questionnaire were distributed, reflecting all possible orderings of the three stimulus scenarios.

The second task asked the same respondents to imagine that they must now choose which proposal among the three submitted should receive the $10 million. They were told that partial funding was not possible and they must rank order the three medical institutions. Before giving their preference order, respondents were told that: (a) Institution X would reduce deaths from Disease A from approximately 15,000 per year to about 5,000 per year, (b) Institution Y would reduce deaths from Disease B from approximately 160,000 per year to about 145,000 per year, and (c) Institution Z would reduce deaths from Disease C from approximately 290,000 per year to about 270,000 per year (see "Task 2 Information" in Table 2).

Thus, there was an inverse relationship between "number of lives saved" and "proportion of lives saved": Disease C deaths were to be reduced by the greatest number (20,000) but by the smallest percentage (7%), whereas Disease A deaths were to be reduced by the smallest number (10,000) but by the greatest percentage (67%). Respondents were asked to rank the three proposals from most worthy to least worthy to receive the $10 million funding.

Results and Discussion

The results from Task 1 indicated that a majority of respondents exhibited psychophysical numbing (see Table 3). When estimating the minimum number of lives an institution's treatment must save to merit a $10 million award, 65% of participants gave estimates that increased as the size of the population at risk increased. Approximately 28% required that the same number be saved, regardless of size, and 7% gave either varied or decreasing estimates. We also calculated the medians and geometric means, which are less affected by extreme values. Table 3 shows that the arithmetic means, medians, and geometric means all reflect a substantial effect consistent with psychophysical numbing.

For those who responded in accord with the numbing hypothesis, the median number of lives Institution Y's treatment was required to save (mdn. = 60,000) was more than six times greater than that of X's treatment (mdn.= 9,000), whereas the median estimate for Z's treatment (mdn. = 100,000) was more than 11 times greater than the estimate for X's treatment. Interestingly, for the 28% of respondents whose estimates did not vary with disease size, the median estimate was only 100, far less than that of the psychophysical numbing respondents' estimates for even the small-scale disease (mdn. = 9,000).

Figure 3 shows the proportion of lives that respondents required each disease treatment to save. Three clearly defined groups emerged from the analysis: 16 respondents (10%) made estimates such that the proportion of lives saved remained constant across disease size; 91 respondents (55%) made estimates such that, as disease size increased, the proportion saved decreased, but at a rate where the number required to be saved was greater for larger diseases; and 47 respondents (28%) made estimates such that, as the disease size increased, the proportion saved decreased at a rate such that the number of lives saved remained constant. The remaining 11 individuals (7%) exhibited no consistent pattern, and were therefore not included in any of the three groups mentioned.

These results suggest that respondents evaluated the interventions using two evaluation strategies. Some respondents appeared to employ a proportion rule; some, an absolute number rule; and still others seemed to employ some combination of the two. That is, some respondents believed a given institution to be "worthy" of funding only if the ratio of number of lives saved to total number of lives at risk reached some proportion threshold (proportion rule). Those who followed the number rule also held a threshold, but one that was based on an absolute number of lives saved rather than a proportion saved. The majority of respondents, however, fell into a third group that made estimates in a manner consistent with an anchoring and adjustment process. In such a process, for example, one might choose a reasonable proportion threshold for the smallest disease treatment, say "must save at least 50% of those at risk", and then adjust this threshold downward for the treatments of larger diseases, say "save 47%" for the medium disease and "save 45%" for the largest disease. In fact, many of the 55% of respondents who appeared in this category imposed proportion thresholds that were within 5% across the three disease treatments.

In Task 2, we gave information about the annual expected reduction in deaths for each of the three disease treatments (see "Task 2 information" in Table 2). Respondents then rank ordered the treatments from "Most worthy" to "Least worthy" to receive funding. The results were quite different from those of Task 1 (see Table 4). More than 60% of respondents preferred to fund disease treatments that maximized the number of lives saved, preferring Institution Z (20,000 saved) over Institution Y (15,000 saved) over Institution X (10,000 saved). Approximately 16% preferred treatments that maximized the proportion of lives saved, choosing the preference order XYZ. The remaining 34% of respondents fell somewhere in between the above two groups, choosing preference orders that maximized neither the number of lives saved nor the proportion of lives saved.

The three groups that emerged in Task 2 roughly corresponded to those that emerged in Task 1, though the proportion of respondents in each did not. In Task 2, a majority of respondents seemed to employ the absolute number rule by consistently preferring to save a greater number of lives; others seemed to employ a proportion rule by consistently preferring to save a greater proportion of lives, and about one third seemed to attempt some combination of the two strategies.

The results of Task 2 underscore two important points. First, the task format can significantly influence the degree of psychophysical numbing. Second, even in the rather simple and transparent format studied here, psychophysical numbing does not disappear.

4. General Discussion

Evidence from the present studies shows that people often judge the value of life-saving efforts in much the same way they judge the intensity of stimuli in traditional psychophysical experiments. Just as a fixed decrease in brightness seems greater when the original intensity is small than when it is large, an intervention saving a fixed number of lives seems more valuable when fewer lives are at risk to begin with—when the savings is a larger proportion of the number at risk. When such psychophysical numbing occurs, the value of a life-saving intervention is inversely proportional to the magnitude of the threat rather than being determined by the absolute number of lives the intervention can save.

A significant portion of the respondents in each of the present studies exhibited psychophysical numbing. Many respondents in the first two studies judged interventions serving larger refugee camps to be considerably less valuable than ones serving smaller camps, even though the interventions could save the same number of lives. Furthermore, when respondents in Task 1 of Study 3 estimated the number of lives several proposed disease treatments would have to save to be equally worthy of a fixed amount of funding, median estimates were more than 11 times greater for the intervention that treated the disease killing 290,000 annually than the one killing 15,000 annually.

Although psychophysical numbing was present in each study, its prevalence varied. This is important because it shows that the incidence of the phenomenon is mutable. For example, over 67% of participants responded psychophysically in Task 1 of Study 3, whereas only 16% responded psychophysically in Task 2. In addition, Study 3 suggested that those most likely to see the value of an intervention as independent of a problem's size were also the ones who attached the greatest value to saving lives generally. For example, psychophysical respondents in Study 3 required over 1,000 times as many lives to be saved (median estimate) in the largest disease category than respondents who gave consistent responses (see Table 3)!

There are several other features in the present studies that may have affected the degree to which psychophysical numbing occurred. First, the way information about life-saving interventions was framed changed the degree of numbing. In Study 3, for example, numbing was frequent when information about the interventions highlighted the magnitude of each tragedy (Task 1), but far less frequent when the information emphasized the magnitude of each intervention's life-saving potential (Task 2). Thus, descriptions of events that focus on the outcomes of the intervention rather than the tragedy it serves appear to reduce the degree of psychophysical numbing.

Second, the ease of comparison between different interventions may have also contributed to the degree of numbing respondents exhibited. In Task 2 of Study 3, where the numbing incidence was low, information on each intervention and tragedy was presented side-by-side, whereas in Task 1, information about each tragedy was presented on separate pages. Study 1 also showed that ease of comparison may have been a factor. Those who, when asked directly, reported no preference between two interventions that saved the same number of lives, nevertheless preferred the intervention serving the smaller tragedy in the previous task where direct comparisons were more difficult.

Despite this variability, however, the present studies suggest that psychophysical numbing is a robust phenomenon—ingrained in the workings of our cognitive and perceptual systems, which seem geared to sensitize us to small changes in our environment, perhaps at the cost of making us less able to appreciate and respond adequately to large changes. When we contemplate nuclear war, for example, and its immense capability for death and destruction, it may be difficult to escape psychophysical numbing as we attempt to grasp the significance of the difference between 10,000, 100,000, or a million or more deaths. Where we lack perceptual sensitivity, we might also expect to find that our language is also inadequate to discriminate among degrees of harm or destructiveness. Thus John Hersey's elegant chronicle of the aftermath of the Hiroshima bombing (which killed about 140,000 people) simply refers to the scene as havoc (Hersey, 1946, p. 5) and terrible (p. 86). Lifton (1967) refers to Hiroshima as a disaster, a term commonly applied to events that are far less severe. Holocaust, catastrophe, calamity, tragedy . . . the vocabulary of disaster seems sparse indeed. Can the potential deaths of large numbers of people really be comprehended without an adequate vocabulary of destructiveness?

Some who have worried about the incomprehension of mass destruction are pessimistic. Humphrey (1981), for example, writes of our ability to be moved greatly by the plight of single human beings at the expense of insensitivity to "giant dangers." He says:

> "In a week when 3,000 people are killed by an earthquake in Iran, a lone boy falls down a well shaft in Italy—and the whole world grieves. Six million Jews are put to death in

Hitler's Germany, and it is Anne Frank trembling in her garret that remains stamped in our memory."

We must live with this . . . It will not change. I do not expect my dog to learn to read The Times, and I do not expect myself or any other human being to learn the meaning of nuclear war or to speak rationally about megadeaths . . . " (pp. 21-22).

Yet writers such as Hersey, Lifton, Jonathan Schell (1982), and many others do have the power to move us emotionally with their eloquent descriptions of individual and societal tragedies despite the lack of adequate one-word descriptors. Is that a sign that we can, indeed, comprehend these tragedies in a way that will help us to make good decisions about preventing them or managing their risks?

Modern technology has great power to cause, prevent, and alleviate mass human suffering. Yet the psychophysical numbing we have observed in our studies is strong enough and pervasive enough to raise some disturbing questions about our ability to make rational decisions when many lives are at stake. Further research is clearly needed to illuminate the dynamics of psychophysical numbing and determine its effects on decision making.

Footnotes

* We acknowledge with thanks support from the National Science Foundation under Grant Number SBR-9422754 to Decision Research and a Graduate Research Fellowship granted to the first author. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation. We also thank Sarah Lichtenstein for her comments on the manuscript and Leisha Mullican for her technical assistance in the preparation of the report.

1. A striking example of psychophysical insensitivity to money was Ronald Reagan's assertion that a $4.6 billion job program "would add virtually nothing to the federal budget" (*Eugene Register-Guard,* March 25, 1983).

2. Lifton (1967) used the term "psychic numbing" to refer to the accommodation and reduced sensitivity to shocking and emotionally overwhelming threats and experiences, such as those created by nuclear war. Hiroshima survivors, for example, said that they very quickly "ceased to feel." We use the term "psychophysical numbing" to reflect a more cognitive or perceptual form of insensitivity as opposed to the affective quality of psychic numbing. Whereas psychic numbing is adaptive, enabling survivors and rescue workers to cope with trauma, psychophysical numbing may degrade our ability to appreciate the consequences of our actions.

3. In the summer of 1994, when these data were collected, ethnic warfare in Rwanda had resulted in over a million refugees fleeing into neighboring Zaire. The brutal nature of the civil war, as well as the problems of disease and hunger that plagued the refugees, had made the Rwandan conflict the topic of considerable print, radio, and television news attention (Cooper, 1994; Purvis, 1994; World News Tonight, 1994a, 1994b). Participants' likely familiarity with the

Rwanda crisis should have helped to make the potential loss of life addressed in the judgment task particularly salient and realistic.

References

"38,000 Shoes Stand For Loss In Lethal Year," (1994) The Register-Guard, September 21, 6A.

Cooper, A. (1994). All Things Considered. Washington, DC: National Public Radio, July 25.

Fechner, G. T. (1860). Elemente der Psychophysik. Leipzig: Breitkopf und Härtel.

Galanter, E. (1962). "The Direct Measurement Of Utility And Subjective Probability," American Journal of Psychology 75, 208-220.

Hersey, John. (1946). Hiroshima. New York: Bantam.

Humphrey, N. (1981). "Four Minutes To Midnight." Quoted in R. J. Lifton and N. Humphrey (eds.), In A Dark Time. Cambridge, MA: Harvard University, 21-22.

Kahneman, Daniel, and Amos Tversky. (1979). "Prospect Theory: An Analysis of Decision Under Risk," Econometrica 47, 263-291.

Lifton, R. J. (1967). Death In Life: Survivors Of Hiroshima. New York: Random House.

Purvis, A. (1994). "All the Hatred in the World," Time, June 13, 36-37.

Rummel, R. J. (1995). "The Holocaust in Comparative and Historical Perspective." Paper delivered at the conference on "The 'Other' as Threat—Demonization and Antisemitism." Hebrew University of Jerusalem, June.

Schell, J. (1982). The fate of the earth. New York: Alfred A. Knopf.

Stevens, S. S. (1975). Psychophysics. New York: Wiley.

Summers, C., P. Slovic, D. Hine, and D. Zuliani. (in press). "'Psychophysical Numbing': An Empirical Basis for Perceptions of Collective Violence." In C. Summers and E. Markusen (eds.), Collective Violence: Harmful Behavior in Groups and Governments.

Tversky, Amos, and Daniel Kahneman. (1981). "The Framing of Decisions and the Psychology of Choice," Science 211, 453-458.

Weber, E. H. (1834). "De Pulsu, Resorptione, Auditu et Tactu." In Annotationes Anatomical et Physiological. Leipzig: Koehler.

World News Tonight. (1994a). "Many Rwandan Refugees Are Dying of Cholera." New York and Washington, DC: American Broadcasting Corporation, July 22.

World News Tonight. (1994b). "U.S. Army Begins Water Purification in Rwanda." New York and Washington, DC: American Broadcasting Corporation, July 26.

Table 1

Summary of Information in the Eight Scenarios Given to All Respondents in Study 2

Scenario number	Zairian refugee camp	Camp size	Water system reliability	Prior aid	Post aid
1	Moga 1	11,000	100%	5%	50%
2	Moga 2	11,000	60%	5%	50%
3	Fizi 1	11,000	100%	50%	95%
4	Fizi 2	11,000	60%	50%	95%
5	Uvira 1	250,000	100%	5%	7%
6	Uvira 2	250,000	60%	5%	7%
7	Kalehe 1	250,000	100%	93%	95%
8	Kalehe 2	250,000	60%	93%	95%

Note. The prior-aid variable indicates the amount of pure water need being met for disease victims in a camp before the aid was delivered. Post aid indicates the water need that would be met for disease victims after the aid was provided. Within each level of plane reliability, the intervention in each camp was capable of keeping the same number of disease victims (1500) alive each day (which usually saves the victims' lives).

Table 2

Information Given to Respondents in Task 1 and Task 2 of Study 3

		Number of deaths per year		
		Task 1	Task 2	
Medical institution	Disease treated	Last year	This year	Next year
X	A	15,000	15,000	5,000
Y	B	160,000	160,000	145,000
Z	C	290,000	290,000	270,000

Note. In Task 1, respondents were asked to indicate the minimum number of lives the treatment would have to save to merit $10 million in funding. In Task 2, respondents were asked to rank order the three programs with regard to priority for receiving $10 million in support.

Table 3

Estimated Minimum Number of Lives Each Institution Would Be Required to Save in Task 1 of Study 3

	Psychophysical numbing respondents n = 107 (65%)			Consistent respondents n = 47 (28%)		
	Institution			Institution		
	X	Y	Z	X	Y	Z
Arithmetic mean	7746	63780	111625	3047	3047	3047
Geometric mean	4701	32678	56707	100	100	100
Median	9000	60000	100000	100	100	100
Lower quartile	3000	16000	29000	1	1	1
Upper quartile	10000	100000	200000	5000	5000	5000

Note: Institutions X, Y, and Z each proposed treatment for a disease that caused 15,000, 160,000, and 290,000 annual deaths, respectively.

Table 4

Task 2 in Study 3: Frequency and Percentage of Respondents' Preference Orders (N = 164)

	Maximize number of lives saved	Maximize proportion of lives saved	Other			
Preference order →	ZYX	XYZ	XZY	YXZ	YZX	ZXY
Frequency	99	26	15	9	7	8
Percentage	60%	16%	9%	5%	4%	5%

Note: Institution X proposed a treatment that would reduce annual deaths by 10,000 (67% of those at risk); Institution Y proposed a treatment that would reduce annual deaths by 15,000 (9%); Institution Z proposed a treatment that would reduce annual deaths by 20,000 (7%).

Figure Captions

Figure 1. Main effects in Study 1 for Rwanda camp size (11,000 or 250,000) and program type (transportation or employment) using preference ratings from paired comparisons. Ratings were coded on a 13-point scale (–6 to +6). Positive numbers indicate preference for a Rwanda program over a non-Rwanda program.

Figure 2. Main effects in Study 2 for the three within-subjects variables: camp size (11,000 or 250,000), prior aid (low or high), and water-system reliability (60% or 100%). Benefit rating was scored on a scale from 0 (low) to 8 (high).

Figure 3. Respondents' estimates in terms of proportion of lives that each institution should save (Study 3).

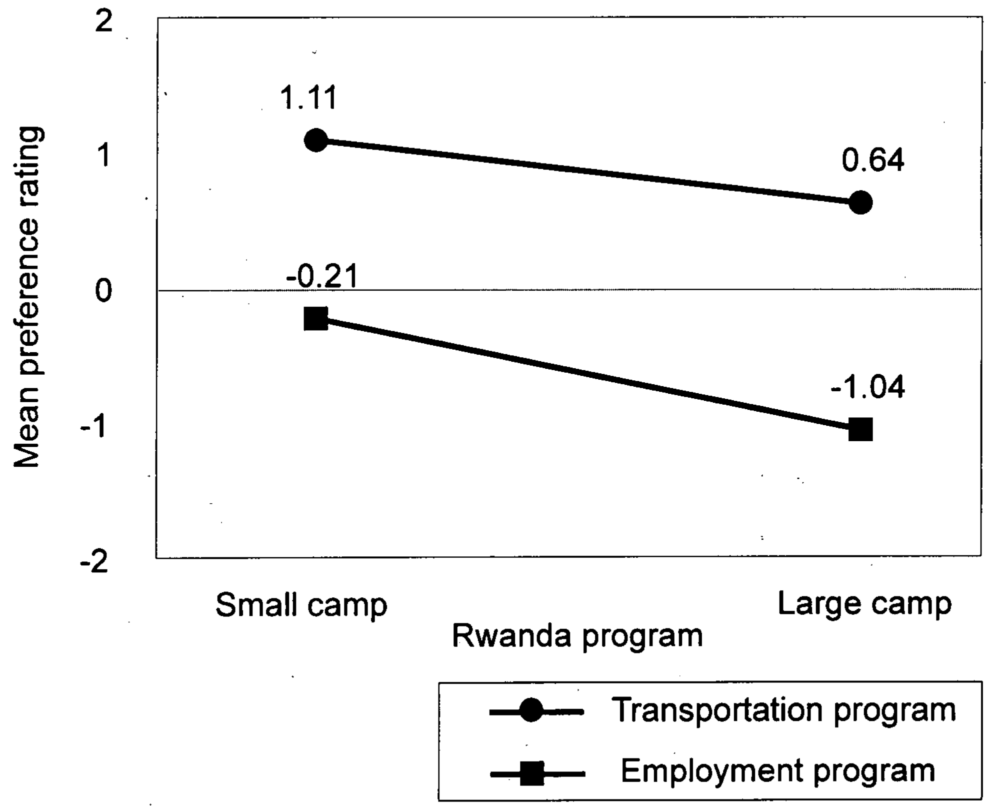

Figure 1. Main effects for Rwanda camp size (11,000 or 250,000) and program type (transportation or employment) using preference ratings from paired comparisons. Ratings were made on a 13-point scale (–6 to +6). Positive numbers indicate preference for a Rwanda program over a non-Rwanda program.

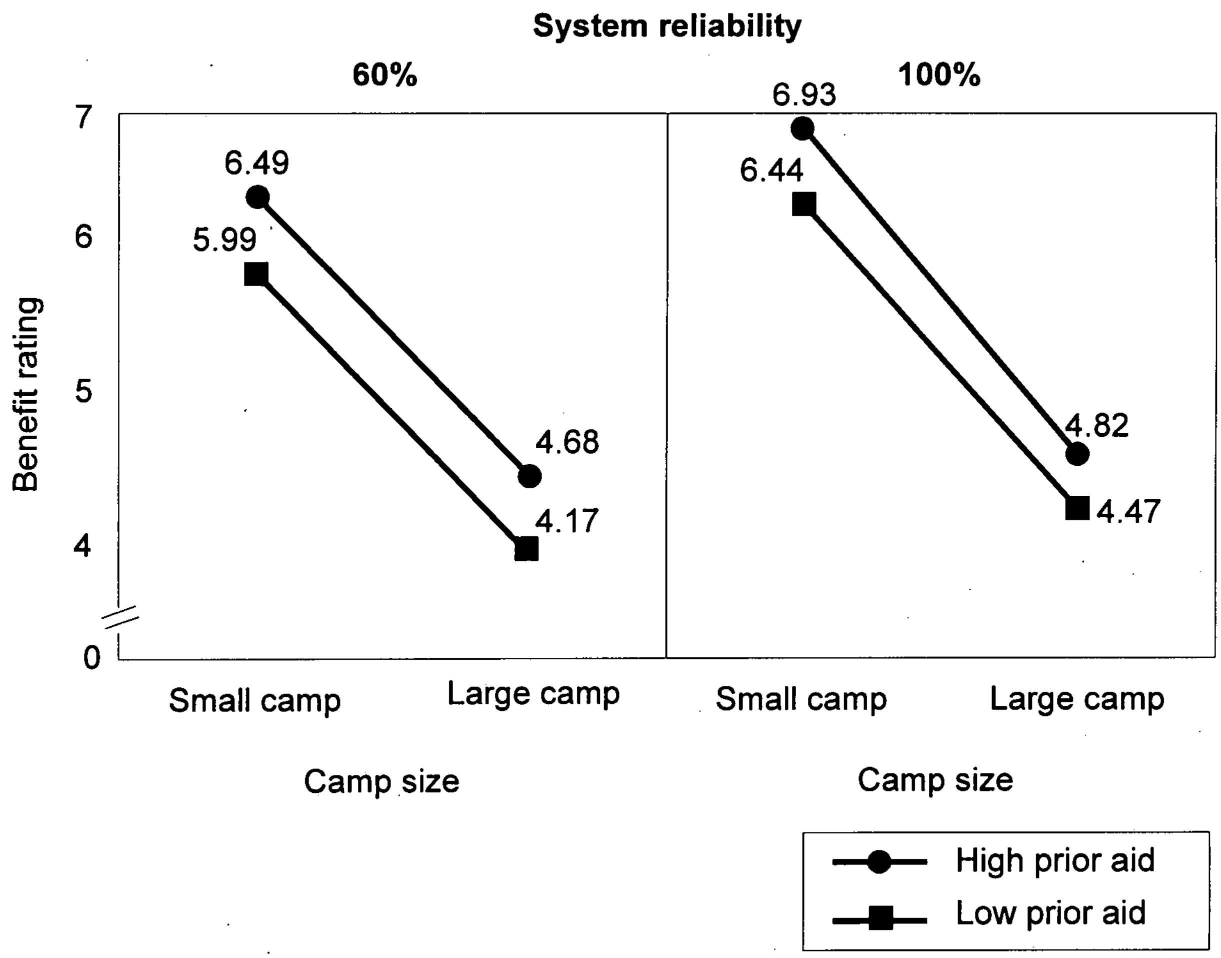

Figure 2. Main effects in Study 2 for the three within-subjects variables: camp size (11,000 or 250,000), prior aid (low or high), and water-system reliability (60% or 100%). Benefit rating was scored on a scale from 0 (low) to 8 (high).

c:\paul\david_f4.cdr 12-23-96

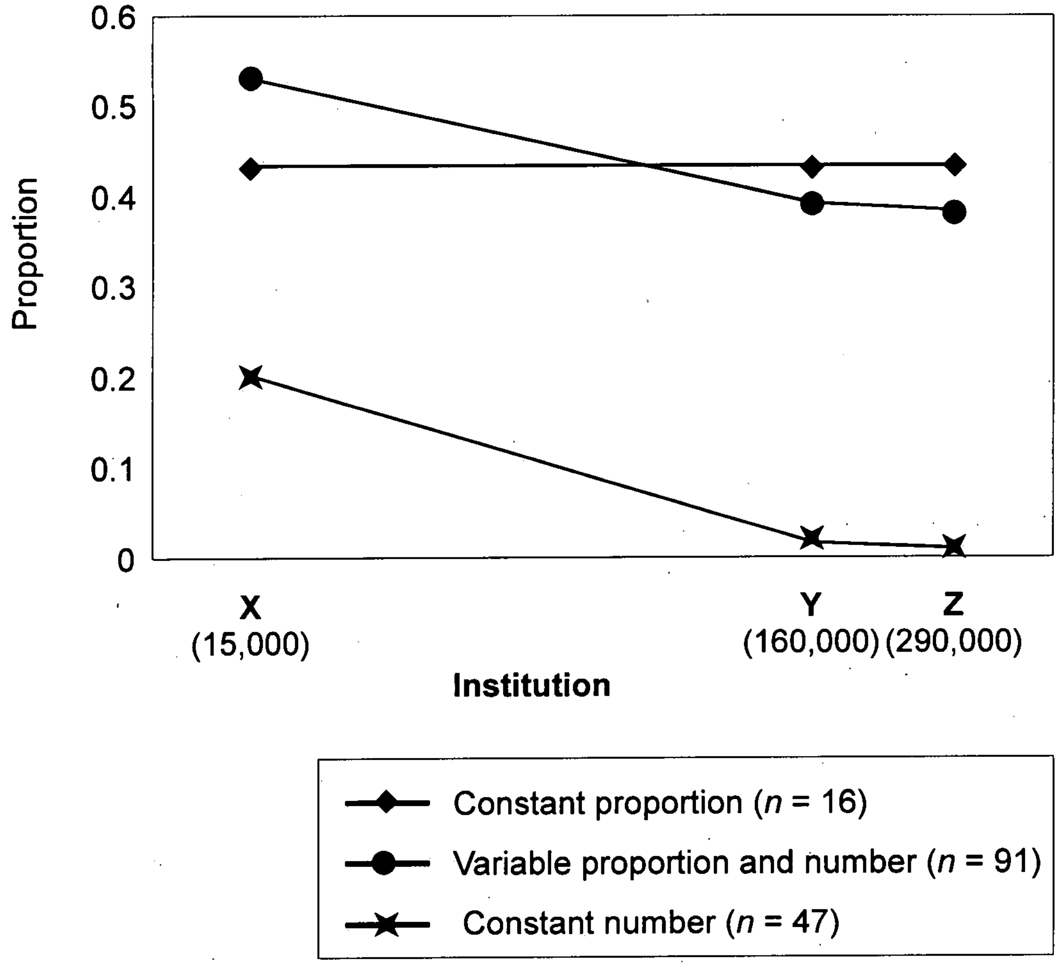

Figure 3. Respondents' estimates in terms of proportion of lives that each institution should save (Study 3).

paul\david_f6.cdr: 12-23-96