Policy Sciences 9 (1978), pp. 127–152

How Safe is Safe Enough? A Psychometric Study of Attitudes Towards Technological Risks and Benefits*

BARUCH FISCHHOFF, PAUL SLOVIC, SARAH LICHTENSTEIN
Decision Research, A Branch of Perceptronics, Eugene, Oregon

STEPHEN READ
University of Texas at Austin

BARBARA COMBS
University of Oregon

ABSTRACT

One of the fundamental questions addressed by risk-benefit analysis is "How safe is safe enough?" Chauncey Starr has proposed that economic data be used to reveal patterns of acceptable risk-benefit tradeoffs. The present study investigates an alternative technique, in which psychometric procedures were used to elicit quantitative judgments of perceived risk, acceptable risk, and perceived benefit for each of 30 activities and technologies. The participants were seventy-six members of the League of Women Voters. The results indicated little systematic relationship between perceived *existing* risks and benefits of the 30 risk items. Current risk levels were generally viewed as unacceptably high. When current risk levels were adjusted to what would be considered acceptable risk levels, however, risk was found to correlate with benefit. Nine descriptive attributes of risk were also studied. These nine attributes seemed to tap two basic dimensions of risk. These dimensions proved to be effective predictors of the tradeoff between acceptable risk and perceived benefit. The limitations of the present study and the relationship between this technique and Starr's technique are discussed, along with the implications of the findings for policy decisions.

* This research was supported by National Science Foundation Grant OEP75-20318 to the University of California, Los Angeles under subcontract No. K559081-0 to Oregon Research Institute. We thank Robyn Dawes, Lewis Goldberg, David Okrent, Ola Svenson, and Chris Whipple for comments on an earlier draft. Correspondence and requests for reprints may be addressed to the authors at Decision Research, 1201 Oak Street, Eugene, Oregon 97401, U.S.A.

Citizens of modern industrial societies are presently learning a harsh and discomforting lesson—that the benefits from technology must be paid for not only with money, but with lives. Whether it be ozone depletion and consequent skin cancer from the use of spray cans, birth defects induced by tranquilizing drugs, or radiation damage from nuclear energy, every technological advance carries some risks of adverse side effects.

Reduction of risk typically entails reduction of benefit, thus posing serious dilemmas for society. With increasing frequency, policy makers are being required to "weigh the benefits against the risks" when making decisions about technological enterprises. To do this, they have been turning to risk-benefit analysis, an offshoot of cost-benefit analysis that is still in its early stages of development, as the basic decision-making methodology for societal risk-taking (Fischhoff, 1977).

The basic question that risk-benefit analysis must answer is: Is this product (activity, technology) acceptably safe? Alternatively, how safe is safe enough?

There are, at present, two main approaches to answering these questions. One, the "revealed preference" method advocated by Starr (1969), is based on the assumption that by trial and error society has arrived at an "essentially optimum" balance between the risks and benefits associated with any activity. One may therefore use economic risk and benefit data from recent years to reveal patterns of acceptable risk-benefit tradeoffs. Acceptable risk for a new technology is defined as that level of safety associated with ongoing activities having similar benefit to society. The present study investigates an alternative approach, called "expressed preferences," which employs questionnaires to measure the public's attitudes towards the risks and benefits from various activities. Both approaches have their proponents and critics (e.g., Kates, 1975; Linnerooth, 1975; Otway and Cohen, 1975).

Starr (1969) illustrated the potential usefulness of revealed preferences by examining the relationship between risk and benefit across a number of common activities. His measure of risk for these hazardous activities was the statistical expectation of fatalities per hour of exposure to the activity. Benefit was assumed to be equal to the average amount of money spent on an activity by an individual participant, or alternatively, equal to the average contribution that the activity makes to a participant's annual income.

From this analysis, Starr derived what might be regarded as "laws of acceptable risk;" namely, that (1) the acceptability of risk is roughly proportional to the third power (cube) of the benefits; (2) the public seems willing to accept risks from voluntary activities (e.g., skiing) roughly 1000 times greater than it would tolerate from involuntary activities (e.g., food preservatives) that provide the same level of benefit; (3) the acceptable level of risk is inversely related to the number of persons exposed to that risk; and (4) the level of risk tolerated for voluntarily accepted hazards is quite similar to the level of risk from disease. On the basis of this last observation, Starr (1969) conjectured that: "The rate of death from disease appears to play, psychologically, a yardstick role in determining the acceptability of risk on a voluntary basis" (p. 1235). Figure 1 depicts the results of Starr's analysis in a *revealed preference* risk-benefit space.

Starr's approach has the advantage of dealing with public behavior rather than with attitudes. It has, however, a number of serious drawbacks. First, it assumes that past behavior is a valid indicator of present preferences. In Starr's words, "The . . . assumption is that historically revealed social preferences and costs are sufficiently enduring to permit their use for predictive purposes" (Starr, 1969, p. 1232). However, Starr and his colleagues have subsequently acknowledged that "The societal value system fluctuates with time, and the technological capability to follow fast-changing societal goals does not exist" (Starr et al., 1976, pp. 635–636). Second Starr's approach "does not serve to distinguish what is 'best' for society from what is 'traditionally acceptable' " (Starr, 1969, p. 1232). What is accepted in the market place may not accurately reflect the public's safety preferences. Consider the automobile, for example. Unless the public really knows what is possible from a design standpoint and unless the automobile industry provides the public with a varied

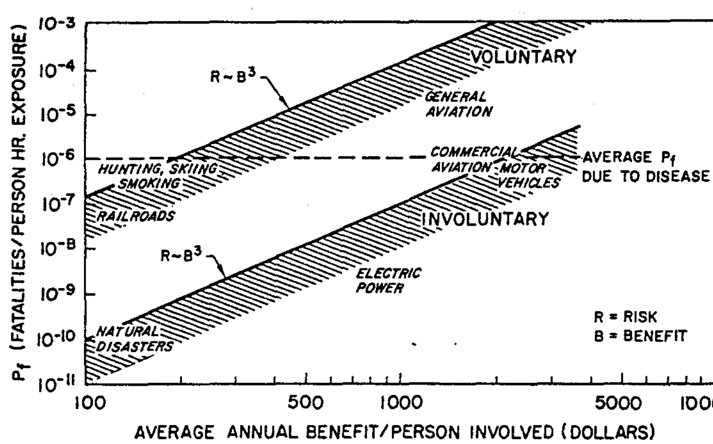

Fig. 1. Revealed risk-benefit relationships (taken from Starr, 1972).

set of alternatives from which to choose, market behavior may not indicate what "a reflective individual would decide after thoughtful and intensive inquiry." A revealed preference approach assumes that people not only have full information, but also can *use* that information optimally, an assumption which seems quite doubtful in the light of much research on the psychology of decision making (Slovic et al., 1977). Finally, from a technical standpoint, Otway and Cohen (1975) have shown that the quantitative conclusions one derives from an analysis of the type Starr performed are extremely sensitive to the way in which measures of risk and benefit are computed from the historical data.

Although only a few questionnaire studies have specifically considered levels of acceptable risk (e.g., Maynard et al., 1976), or the value of a life at risk (Acton, 1973; Torrance, 1970), direct questioning procedures have been used to scale the perceived seriousness of a wide variety of natural and man-made hazards (see, for example, Wyler, Masuda and Holmes, 1968; Golant and Burton, 1969; Otway, Maderthaner & Gutmann, 1975; Otway & Pahner, 1976; Lichtenstein et al., in press).

Use of psychometric questionnaires has been criticized on the grounds that answers to hypothetical questions bear little relationship to actual behavior.

> Time and time again, action has been found to contradict assertion. Since surveys always elicit some degree of strategic behavior ('What do they want me to say?'), we would be better advised to observe what people choose under actual conditions (Rappaport, 1974, p. 4).

Such criticisms of psychometric studies appear to us to be overstated. Attitudes elicited in surveys often correlate highly with behavior (Liska, 1975). Furthermore, they elicit present values rather than historical preferences.

The goal of the present study is to evaluate the usefulness of questionnaire techniques for investigating issues pertaining to risk-benefit tradeoffs. Psychometric procedures were used to elicit quantitative judgments of perceived risk and benefit from various activities and technologies as well as judgments of acceptable risk levels. Participants in our experiment also judged the degree of voluntariness of each activity or technology. These judgments were used to determine whether people do, indeed, judge the acceptability of risks differently for voluntary and involuntary activities. The influence of other potential moderators of perceived and acceptable risk were also studied. These included familiarity with the risk, its perceived controllability, its potential for catastrophic (multiple-fatality) consequences, the immediacy of its consequences, and the extent of scientists' and the public's knowledge about its consequences. Various authors have speculated about the influence of these factors (e.g., Green, 1974; Lowrance, 1976; Otway, 1975; Otway & Pahner, 1976; Rowe, 1977; Starr et al., 1976), but little empirical data is available.

Method

Design

The participants in our study evaluated each of 30 different activities and technologies with regard to (1) its perceived benefit to society; (2) its perceived risk, (3) the acceptability of its current level of risk; and (4) its position on each of nine dimensions of risk. As tasks (1) and (2) were quite arduous and as we were interested in independent judgments of perceived risk and benefit, participants performed either tasks (1), (3) and (4) or tasks (2), (3) and (4). Which of the two combinations of tasks they faced was determined randomly. As part of their general instructions, participants were told, "This is a difficult, if not impossible, task. Nevertheless, it is not unlike the task you face when you vote on legislation pertaining to nuclear power, handguns, or highway safety. One never has all the relevant information; ambiguities and uncertainties abound, yet some judgment must be made. The present task should be approached in the same spirit."

Items

The 30 activities and technologies included the eight items used by Starr (1969) and 22 others chosen to vary broadly in the quality and quantity of their associated risks and benefits. They appear in Table 1.

Tasks

Perceived Benefit

People given this task were asked to "consider all types of benefits: how many jobs are created, how much money is generated directly or indirectly (e.g., for swimming, consider the manufacture and sale of swimsuits), how much enjoyment is brought to people, how much of a contribution is made to the people's health and welfare, and so on." Thus, they were told to give a global estimate of all benefits, both tangible and intangible. They were specifically told: "Do not consider the costs of risks associated with these items. It is true, for example, that swimmers sometimes drown. But evaluating such risks and costs is not your present job. Your job is to assess the *gross benefits,* not the net benefits which remain after the costs and risks are subtracted out. Remember that a beneficial activity affecting few people will have less gross benefit than a beneficial activity affecting many people. If you need to think of a time period during which the benefits accrue, think of a whole year—the total value to society from each item during one year."

In order to make the evaluation task as easy as possible, each activity appeared on a 3 × 5 inch card. Participants were told first to study the items individually, thinking of the benefits accruing from each; then to order them from least to most beneficial; and finally, to assign numerical benefit values by giving a rating of 10 to the least beneficial and making the other ratings accordingly. They were also given additional suggestions, clarifications and encouragement to do as accurate a job as possible. For example, they were told "a rating of 12 indicates that the item is 1.2 times as beneficial as the least beneficial item (i.e., 20% more beneficial). A rating of 200 means that the item is 20 times as beneficial as the least beneficial item, to which you assigned a 10 . . . Double-check your ratings to make certain that they are consistent. For example, if one activity is rated 50 and a second 100, the second item should seem twice as beneficial as the first. Adjust the numbers until you feel that they are right for you."

Perceived Risk

Participants in this task (who, it will be remembered, did not judge perceived benefit) were told to "consider the risk of dying as a consequence of this activity or technology. For example, use of electricity carries the risk of electrocution. It also entails risk for miners who produce the coal that generates electricity. Motor vehicles entail risk for drivers, passengers, bicyclists and pedestrians, etc." They were asked to order and rate these activities for risk with instructions that paralleled the instructions for the perceived benefit task, giving a rating of 10 to the least risky item and scaling the other items accordingly.

Note

These measures of risk and benefit differ from Starr's in several respects other than their source in attitudes rather than in behavior. Our subjects were asked to evaluate total risk per year to participants, not risk per hour of exposure, the unit of measurement used by Starr (1969). Several considerations motivated this change of unit, the most important of which is that the definition of "hour of exposure" is extremely equivocal for some items (e.g., handguns, pesticides). Excluding activities and technologies for which such measurement is problematic would introduce a systematic bias into our sample of items. Although the shape or magnitude of the relationship may vary with choice of measure, there is no a priori reason why people's historical risk-benefit tradeoffs are best revealed with one particular measure of risk. Starr, himself (1972, p. 28), apparently believed that total risk per year was a more appropriate measure but rejected it because of measurement difficulties and because of his belief that "the hour of exposure unit [is] more closely related to the individual's intuitive process in choosing an activity than a year of exposure would be." In any case, he found that use of either unit "gave substantially similar results."

A second difference in unit is that we have considered *total* benefit and risk to society rather than *average* benefit and risk per person involved. For activities and technologies whose risks and/or benefits are shared by all members of society, this change is inconsequential. For the others, our risk and benefit measures should be weighted by the proportion of individuals participating in the activity in order to achieve strict comparability with Starr's measures.

A third difference from Starr is that relying on our participants' ability to consider all types of benefits relieved us of the restriction which Starr imposed upon himself to consider only benefits to which a dollar value could be readily assigned.

Risk Adjustment Factor

After rating risks or benefits, both groups of participants were asked to judge the acceptability of the level of risk currently associated with each item. The instructions included the following:

> This is not the ideal risk. Ideally, the risks should be zero. The acceptable level is a level which is 'good enough,' where 'good enough' means you think that the advantages of increased safety are not worth the costs of reducing risk by restricting or otherwise altering the activity. For example, we can make drugs 'safer' by restricting their potency; cars can be made safer at a cost, by improving their construction or requiring regular safety inspection; we may, or may not, feel restrictions are necessary.
>
> If an activity's present level of risk is acceptable, no special action need be taken to increase its safety. If its riskiness is unacceptably high, *serious action*, such as legislation to restrict its practice, should be taken. On the other hand, there may be some activities or technologies that you believe are currently safer than the acceptable level of risk. For these activities, the risk of death could be higher than it is now before society would have to take serious action.

On their answer sheets, participants were provided with three columns labelled: (a) "Could be riskier: it would be acceptable if it were — times riskier;" (b) "It is presently acceptable;" and (c) "Too risky: to be acceptable, it would have to be — times safer." These risk adjustment factors were used to establish levels of acceptable risk.

Rating Scales

As their final task, participants were asked to rate each activity or technology on nine seven-point scales, each of which represented a dimension which has been hypothesized to influence perceptions of actual or acceptable risk (e.g., Lowrance, 1976). These scales, in the order and wording in which they were described, were:

1. Voluntariness of risk: Do people get into these risky situations voluntarily? If for a single item some of the risks are voluntarily undertaken and some are not, mark an appropriate spot towards the center of the scale. (The scale was labelled: 1 = voluntary; 7 = involuntary.)

2. Immediacy of effect: To what extent is the risk of death immediate—or is death likely to occur at some later time? (1 = immediate; 7 = delayed.)

3. Knowledge about risk: To what extent are the risks known precisely by the persons who are exposed to those risks? (1 = known precisely; 7 = not known.)

4. Knowledge about risk: To what extent are the risks known to science? (1 = known precisely; 7 = not known.)

5. Control over risk: If you are exposed to the risk of each activity or technology, to what extent can you, by personal skill or diligence, avoid death while engaging in the activity? (1 = uncontrollable; 7 = controllable.)

6. Newness: Are these risks new, novel ones or old, familiar ones? (1 = new; 7 = old.)

7. Chronic-catastrophic: Is this a risk that kills people one at a time (chronic risk) or a risk that kills large numbers of people at once (catastrophic risk)? (1 = chronic; 7 = catastrophic.)

8. Common-dread: Is this a risk that people have learned to live with and can think about reasonably calmly, or is it one that people have great dread for—on the level of a gut reaction? (1 = common; 7 = dread.)

9. Severity of consequences: When the risk from the activity is realized in the form of a mishap or illness, how likely is it that the consequence will be fatal? (1 = certain not to be fatal; 7 = certain to be fatal.) Green (1974) has referred to this as the "sporting chance" factor.

Participants rated all 30 activities and technologies on each scale before proceeding to the next.

Participants

Members of the Eugene, Oregon, League of Women Voters and their spouses were asked to participate in the study in return for a contribution to the organization's treasury. In all, 76 individuals (52 women and 24 men) returned completed, anonymous questionnaires. Spouses received the same set of questionnaires and were instructed not to discuss the tasks until they were completed. They indicated having spent an average of two hours on the three tasks. Although League members and spouses are by no means representative of all American adults, they do constitute an extremely thoughtful, articulate, and influential group of private citizens. If there are systematic relationships between people's judgments of risk and benefit, they should be found in these participants' responses. While the particular relationships found here might differ from those found with other populations, the opinions of League members may be quite similar to those of many of the private citizens most heavily engaged in the public policy-making process.

Results

Perceived Risk and Benefit

Because arithmetic means tend to be unduly influenced by occasional extreme values, geometric means were used to describe the data. They are calculated by taking the log of each score, finding the arithmetic mean of those logs and then finding the antilog of the arithmetic mean. Columns 1 and 2 of Table 1 present the geometric means of all risk and benefit judgments for each item.

Many of the substantive results in this table appear to be accurate reflections of the attitudes of a generally liberal, environmentally-minded group, the League of Women Voters. Especially interesting are the low benefit attributed to food coloring, spray cans, and handguns, and the great difference between the evaluations of non-nuclear and nuclear electric power. Although these specific judgments are quite revealing, the main purpose of this study was not to poll the attitudes of any particular group of citizens, but to examine the relationships between perceived benefit and risk.

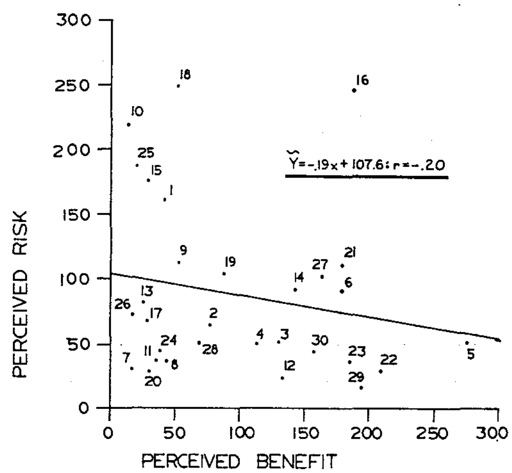

Fig. 2. Relationship between perceived risk and perceived benefit.

Figure 2 presents these judgments in a perceived risk-benefit space analogous to Starr's revealed risk-benefit space (Fig. 1). In general, perceived risk *declined* slightly with overall benefit, motor vehicles being the only item which rates high on both scales. The overall best-fit line had a negative slope ($\hat{y} = -0.19x + 107.6$; $r = -0.20$; $p > 0.25$). The axes in Starr's Fig. 1 are logarithmic. Replotting our data in Fig. 2 using log geometric means left the relationship unchanged ($y = -0.18x + 2.18$; $r = -0.23$).

Examination of columns 1 and 2 in Table 1 provides insight into the nature of the perceived risk-benefit relationship. Society presently tolerates a number of activities that our participants rated as having very low benefit and very high risk (e.g., alcoholic beverages, handguns, motorcycles, smoking) as well as a number of activities perceived

TABLE 1

Mean Judgments of Risk and Benefit from 30 Activities and Technologies

Activity or technology	Perceived benefit (Geometric mean)	Perceived risk (Geometric mean)	Risk adjustment factor [a]: Risk subjects (Geometric mean)	Risk adjustment factor [a]: Benefit subjects (Geometric mean)	Acceptable level of risk [b]: Risk subjects	Acceptable level of risk [b]: Benefit subjects
1. Alcoholic beverages	41	161	4.7	4.2	34	38
2. Bicycles	82	65	1.6	1.4	41	46
3. Commercial aviation	130	52	1.2	1.4	43	37
4. Contraceptives	113	50	2.1	1.9	24	26
5. Electric power [c]	274	52	1.2	0.9	43	58
6. Fire fighting	178	92	1.2	1.0	77	92
7. Food coloring	16	31	2.7	3.4	11	9
8. Food preservatives	44	36	2.6	2.8	14	13
9. General (private) aviation	53	114	2.3	1.8	50	63
10. Handguns	14	220	17.1	17.5	13	13
11. High school and college football	35	37	1.8	1.6	21	23
12. Home appliances	133	25	1.1	1.0	23	25
13. Hunting	30	82	2.9	2.1	28	39
14. Large construction (dams, bridges, etc.)	142	91	2.0	1.4	46	65
15. Motorcycles	29	176	5.1	5.5	35	32
16. Motor vehicles	187	247	7.3	4.9	34	50
17. Mountain climbing	28	68	1.1	0.9	62	76
18. Nuclear power	52	250	32.2	25.9	8	10
19. Pesticides	87	105	10.5	8.5	10	12
20. Power mowers	30	29	1.7	1.3	17	22
21. Police work	178	111	2.3	1.3	48	85
22. Prescription antibiotics	209	30	1.4	1.2	21	25
23. Railroads	185	37	1.4	1.1	26	34
24. Skiing	38	45	1.1	1.0	41	45
25. Smoking	20	189	15.2	15.3	12	12
26. Spray cans	17	73	8.6	6.9	8	11
27. Surgery	164	104	2.2	1.6	47	65
28. Swimming	68	52	1.2	0.9	43	58
29. Vaccinations	194	17	1.0	0.7	17	24
30. X-rays	156	45	2.1	1.3	21	35
All responses	69	69	2.7	2.2		
Coefficient of Concordance	0.77	0.50	0.50	0.50		

[a] Values greater than one mean that the item should be safer; values less than one mean that the item could be riskier.

[b] Acceptable levels of risk were calculated by dividing column 2 by columns 3 and 4 respectively.

[c] Non-nuclear.

to have great benefit and relatively low risk (e.g., prescription antibiotics, railroads, vaccinations).

Could these differences between our results and Starr's be artifacts of technical differences between our research procedures? In particular, are they due to our use of additional technologies and different units of measurement? We will consider these factors in turn. We can, however, make no statement about the degree to which our results depend on the participant population studied.

Different Items

Figure 3 compares the perceived (bottom figure) and revealed (top figure) risk-benefit spaces for the 8 of our 30 items used also by Starr (1969). The computed space is Otway and Cohen's (1975) recalculation of Starr's original data. Although the scale differences make it difficult to compare the two figures directly, it is clear that both

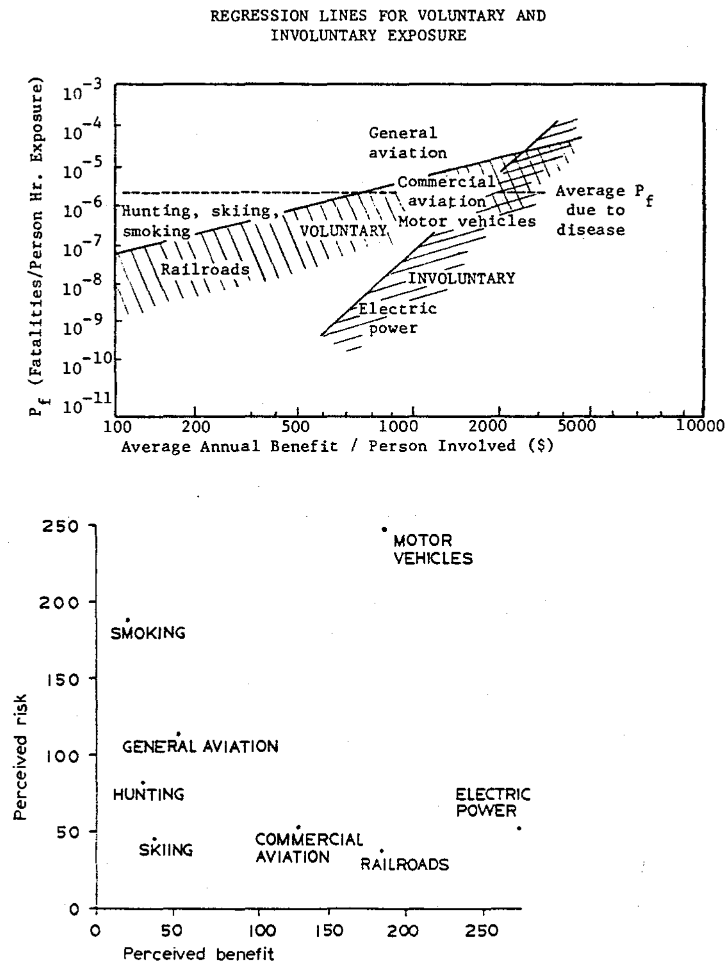

Fig. 3. Relationship between perceived risk and perceived benefit for the items studied by Starr (1969) and Otway and Cohen (1975) (above) and the present subjects (below).

the nature of the relationship and the relative costs and benefits of the various items were quite different. For this subset, as for the full set of items, risk decreased somewhat with benefit in the perceived space.

Different Units

Allowing people to consider all benefits accruing from an activity, not just those readily expressed in dollars, may have been responsible for some of the differences between our results and Starr's. For example, railroads and electric power appear to be relatively more beneficial in the perceived space, perhaps reflecting the not-readily-quantifiable environmental benefits of the former and the "great flexibility in patterns of living" (Starr, 1972, p. 29) conferred by the latter. Here, we believe that there are advantages to using the comprehensive measure of benefit.

Another difference between our units of measurement and Starr's was that we considered total risk and benefit to society, not just consequences per person exposed. Twenty-five of the 30 activities and technologies used in this study have risks and/or benefits for all or almost all members of society. For these, use of risk and benefit per participant (Starr's measure) would produce a figure whose pattern is identical to that in Fig. 2 ($\tilde{y} = -0.20x + 109.7$; $r = -0.21$ without those items, compared to $\tilde{y} = -0.19x + 107.6$; $r = -0.20$ with them).

Finally, we have argued that the unit "risk per year" exposure used here is equal or superior to the unit "risk per hour" exposure used by Starr. Whether this change of unit was responsible for differences in our results is a topic for future research. Starr reported that the change made little difference in his computations, although that need not also be the case with subjective estimates.

Risk Adjustment Factor

Columns 3 and 4 of Table 1 present the geometric means of our participants' judgments of the acceptability of the risk levels associated with the various items. As indicated by the preponderance of items for which the mean adjustment factor is greater than one, people thought that most items should be made safer; this occurred despite instructions emphasizing that such a rating indicated the need for serious societal action. Of the 2280 acceptability judgments, roughly half indicated that the item in question was too risky; 40% indicated that its current risk level was appropriate and 10% indicated that it could be riskier still. There were, however, relatively few items which people believed should be made *much* safer, namely alcoholic beverages, handguns, motorcycles, motor vehicles, nuclear power, pesticides, smoking, and spray cans.

Perceived risk was correlated 0.75 and 0.66 with risk adjustment factor ratings for the risk and benefit groups, respectively. Thus, both groups felt that the higher the risk, the more it should be reduced.

Participants in our study made these risk adjustment ratings after ordering and rating the 30 items for either perceived benefits or perceived risks. Comparing columns 3 and 4 shows that for 24 of the 30 items, the current risk level was judged more acceptable (less in need of change) by those people who had previously considered benefits than by those who had previously dwelt on risks. Thus, the way in which activities and technologies are considered may affect the acceptability of their risk levels.

A "level of acceptable risk" was determined for each item by dividing its perceived risk (column 2 of Table 1) by the geometric mean adjustment factor (column 3 or 4 of Table 1). This was done separately for people who had previously judged risk first and those who had judged benefit first. The results are shown in columns 5 and 6 of Table 1. For example, for alcoholic beverages, the level of acceptable risk was $161/4.7 = 34.2$ for perceived risk participants and $161/4.2 = 38.3$ for perceived benefit participants.

Figure 4 compares each item's acceptable risk level with its perceived benefit. It shows what societal risk-benefit tradeoffs would be if current risk levels were adjusted to acceptable levels. In this figure, the level of acceptable risk increased with the level of perceived benefit, although the relationship was not strong. According to this inferred relationship, participants in our study believed that more risk should be tolerated with more beneficial activities.

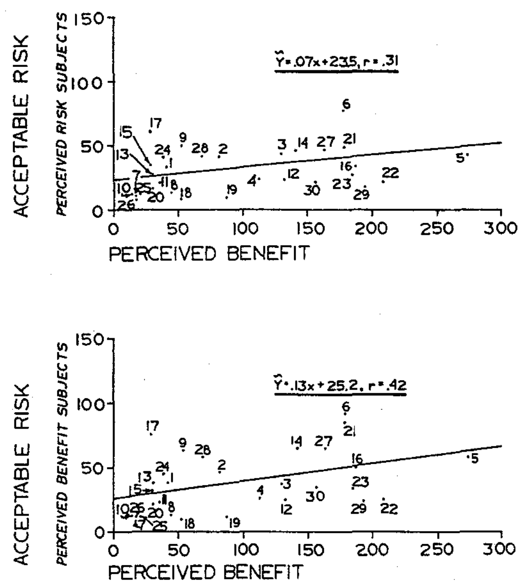

Fig. 4. Relationship between perceived benefit and acceptable risk.

Interparticipant Agreement

In order to assess how well participants agreed with one another in their judgments, Kendall's coefficient of concordance was computed for each task. This index reflects the average rank-order correlation between the judgments of all pairs of participants (Siegel, 1956). The values shown at the bottom of Table 1 are moderate to high (all being significantly different from zero at $p<0.001$) and indicate that people substantially agreed in their rankings, particularly when they evaluated benefits (column 1).

Rating Scales

Table 2 presents the arithmetic mean ratings of the nine risk characteristics for the 30 items. Since people in the perceived risk and perceived benefit groups produced very similar judgments (difference in means less than 1.00 in almost every case), their responses were pooled. There was also considerable agreement among people within each of the two groups, as reflected by the moderately high coefficients of concordance.

TABLE 2

Mean Ratings for Nine Characteristics of Risk

		Voluntariness 1 = voluntary	Immediacy 1 = immediate	Known to exposed 1 = precisely	Known to science 1 = precisely	Controllability 1 = can't be controlled
1. Alcoholic beverages		2.10	5.34	3.77	1.98	5.57
2. Bicycles		1.90	2.82	3.27	2.80	4.99
3. Commercial aviation		2.80	1.85	3.24	2.12	2.18
4. Contraceptives		2.74	5.69	4.66	3.88	3.11
5. Electric power (non-nuclear)		4.40	2.82	3.98	2.68	4.25
6. Fire fighting		2.40	2.33	1.98	2.25	4.03
7. Food coloring		5.86	6.26	6.40	4.77	2.70
8. Food preservatives		5.65	6.18	6.39	4.76	2.70
9. General aviation		2.20	1.66	2.96	2.60	3.99
10. Handguns		3.42	1.65	2.64	2.41	4.05
11. H. S. and college football		1.90	3.52	3.66	3.11	4.15
12. Home appliances		3.61	2.97	4.47	2.90	4.85
13. Hunting		2.01	1.66	2.62	2.64	4.45
14. Large construction		3.07	2.23	2.77	2.51	3.91
15. Motorcycles		1.87	1.76	2.69	2.17	4.08
16. Motor vehicles		4.04	2.33	3.14	2.31	4.19
17. Mountain climbing		1.15	1.78	1.83	2.49	4.98
18. Nuclear power		6.51	5.08	5.85	4.83	1.36
19. Pesticides		5.77	5.57	5.50	4.41	2.14
20. Power mowers		2.23	2.99	3.31	2.60	5.13
21. Police work		2.44	2.14	2.05	2.25	3.76
22. Prescription antibiotics		4.44	4.33	5.40	3.91	2.77
23. Railroads		3.42	2.91	3.66	2.68	3.22
24. Skiing		1.28	2.45	2.47	2.51	4.73
25. Smoking		1.85	6.11	2.86	2.15	4.43
26. Spray cans		3.80	6.06	5.43	4.16	3.60
27. Surgery		4.28	2.71	3.84	2.86	2.39
28. Swimming		1.64	1.76	2.87	2.68	5.17
29. Vaccinations		3.82	3.71	4.84	2.82	2.53
30. X-rays		4.38	6.15	5.05	3.28	2.37
	$\bar{X}$	3.24	3.49	3.78	2.98	3.73
	σ	1.47	1.70	1.33	0.87	1.10
Coefficient of	Benefit Ss	0.59	0.59	0.53	0.30	0.40
Concordance	Risk Ss	0.62	0.66	0.61	0.35	0.42

TABLE 2

Mean Ratings for Nine Characteristics of Risk

	Newness 1 = new	Chronic–Catastrophic 1 = chronic	Common–Dread 1 = common	Severity of consequences 1 = certain not to be fatal
1. Alcoholic beverages	6.61	1.79	1.92	4.40
2. Bicycles	5.19	1.30	1.74	3.77
3. Commercial aviation	4.24	6.09	3.39	5.72
4. Contraceptives	2.25	1.49	3.14	4.08
5. Electric power	5.09	2.66	1.72	4.52
6. Fire fighting	6.01	2.84	2.62	4.42
7. Food coloring	2.66	2.82	3.24	3.59
8. Food preservatives	2.73	2.82	3.32	3.66
9. General aviation	4.08	3.40	3.15	5.63
10. Handguns	5.69	2.10	4.40	5.67
11. H. S. and College football	4.78	1.40	1.95	3.15
12. Home appliances	4.39	1.38	1.43	3.08
13. Hunting	6.14	1.59	2.79	4.91
14. Large construction	5.04	3.04	2.61	4.77
15. Motorcycles	4.31	1.59	3.02	5.19
16. Motor vehicles	4.73	3.28	3.04	4.57
17. Mountain climbing	5.63	1.32	2.57	4.80
18. Nuclear power	1.35	6.43	6.42	5.98
19. Pesticides	2.22	4.75	5.21	4.87
20. Power mowers	3.70	1.16	1.75	2.75
21. Police work	5.50	2.07	3.05	4.35
22. Prescription antibiotics	2.87	2.35	2.19	3.82
23. Railroads	5.49	4.49	1.75	3.60
24. Skiing	4.69	1.06	1.92	3.15
25. Smoking	5.04	1.68	2.89	5.01
26. Spray cans	1.89	3.82	3.62	4.27
27. Surgery	4.95	1.14	4.04	4.68
28. Swimming	6.50	1.16	1.89	4.78
29. Vaccinations	4.50	1.88	2.03	3.62
30. X-rays	4.02	1.99	2.58	4.20
$\overline{X}$	4.41	2.50	2.85	4.37
σ	1.41	1.43	1.12	0.85
Coefficient of Concordance — Benefit Ss	0.59	0.58	0.45	0.45
Coefficient of Concordance — Risk Ss	0.65	0.57	0.47	0.50

Perceived risk and benefit

Correlations between the nine risk characteristics and perceived risk and benefit are shown in Table 3. None of the risk characteristics correlated significantly with perceived benefit (column 1). Perceived risk was found to correlate with dread and severity but not with any of the other characteristics.

Starr (1969) hypothesized that the tradeoff between risk and benefit is mediated by degree of voluntariness. If so, we would expect a tendency for voluntary activities

TABLE 3

Correlations Between Rating Scales and Perceived Risk and Benefit

Scale	Perceived benefit	Perceived risk	Deviations from perceived benefit–perceived risk regression line [a]
Voluntariness (1 = voluntary)	0.24	0.08	0.13
Immediacy (1 = immediate)	—0.15	—0.07	—0.11
Known to exposed (1 = known precisely)	0.04	—0.20	—0.22
Known to science (1 = known precisely)	—0.16	—0.17	—0.21
Controllability (1 = uncontrollable)	—0.29	—0.04	—0.06
Newness (1 = new)	0.14	0.05	0.08
Chronic (1 = chronic)	0.12	0.30	0.29
Common/dread (1 = common)	—0.26	0.64*	0.54*
Severity of consequences (1 = certain not to be fatal)	—0.10	0.67*	0.66*

[a] as shown in Fig. 2.
* $p<0.001$.

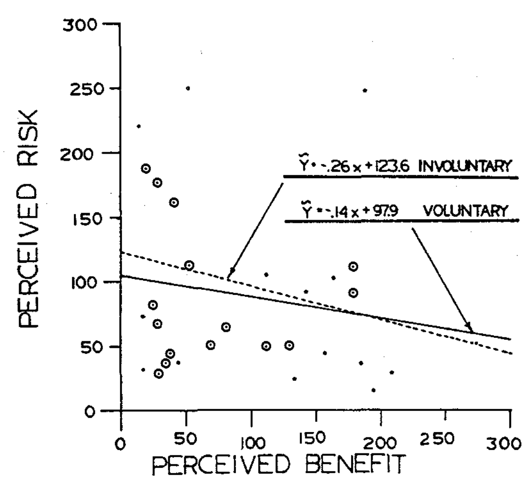

Fig. 5. Relationship between perceived risk and perceived benefit for voluntary (circled) and involuntary hazards.

to lie above the common regression line shown in Fig. 2 and involuntary activities to lie below it. This was not the case. In Fig. 5, the 30 items are dichotomized into the 15 most and 15 least voluntary. Regression lines for the voluntary and involuntary subsets were virtually identical.

If voluntariness is conceptualized as a continuous rather than a dichotomous variable, then according to Starr's hypothesis an item lying high above the common regression line should be very voluntary (have a rating near 1); an item lying far below that line should be very involuntary (have a rating near 7). Let us define a deviation score as the signed vertical distance between each point in the risk-benefit space and the regression line; positive deviation scores belong to points above the line. These deviation scores reflect the variance in the risk scores that cannot be accounted for by the benefit scores. The correlation between these deviation scores and the voluntariness ratings indicates the proportion of this unexplained variance which can be accounted for by the voluntariness measure. Starr's hypothesis suggests that this correlation would be negative (high positive deviations going with low voluntariness ratings). As can be seen from column 3 of Table 3, this was not the case ($r = 0.13$).

However, two other risk characteristics, commonness and severity, did correlate with the deviations from the regression line. If we drew a figure for each of these scales like Fig. 5, we would find two roughly parallel regression lines, one lying above the other. With the scale common/dread, we would find that items whose

TABLE 4

Correlation Between Rating Scales and Measures of Acceptable Risk (Perceived risk and perceived benefit groups combined)

Scale	Risk Adjustment factor	Level of acceptable risk [a]	Deviations from perceived benefit-level of acceptable risk regression line [b]
Voluntariness (1 = voluntary)	0.38*	–0.47†	–0.64‡
Immediacy (1 = immediate)	0.28	–0.64‡	–0.64‡
Known to exposed (1 = known precisely)	0.21	–0.68‡	–0.75‡
Known to science (1 = known precisely)	0.29	–0.57‡	–0.58‡
Controllability (1 = uncontrollable)	–0.30	0.40*	0.48†
Newness (1 = new)	–0.34	0.60‡	0.60‡
Chronic (1 = chronic)	0.45*	0.22	–0.25
Common/dread (1 = common)	0.75‡	–0.29	–0.24
Severity of consequences (1 = certain not to be fatal)	0.54‡	0.17	0.22

* $p<0.05$. † $p<0.01$. ‡ $p<0.001$.
[a] Perceived risk divided by risk adjustment factor.
[b] As shown in Fig. 3.

consequences are more dread tend to have higher perceived risk, at all levels of benefit, than items with more common consequences. Similarly, the line for the more severe (certain to be fatal) risks would lie above the line for less severe risks.

Acceptable Risk

The fact that *perceived* risk was unrelated to voluntariness does not necessarily contradict Starr's claim that the voluntary nature of an activity influences its *acceptable* risk level. Table 4 presents the correlations between each risk characteristic and various aspects of acceptability. The significant correlations in the first column indicate that activities with the most dread and certainly fatal consequences were deemed most in need of risk reduction. The significant correlations in the second column show that if risks were adjusted to an acceptable level, then higher risk levels would be tolerated for old, voluntary activities with well known and immediate consequences. The correlations in the third column show the extent to which each qualitative risk characteristic accounts for variance in acceptable risk unexplained by perceived benefit. These correlations were significant for each of the first six characteristics. Thus, for any given level of benefit, greater risk was tolerated if that risk was voluntary, immediate, known precisely, controllable, and familiar.

The relationship between voluntariness and acceptable risk level is further illustrated in Fig. 6 which shows the separate regression lines for the 15 most and least voluntary activities and technologies. Figure 6 clearly shows a double standard in risk tolerance for voluntary and involuntary activities, like that found by Starr in Fig. 1. Note, however, that the *y*-axis in Fig. 6 is "acceptable risk level" and not "current risk level". The participants in our study believed that a double standard would be

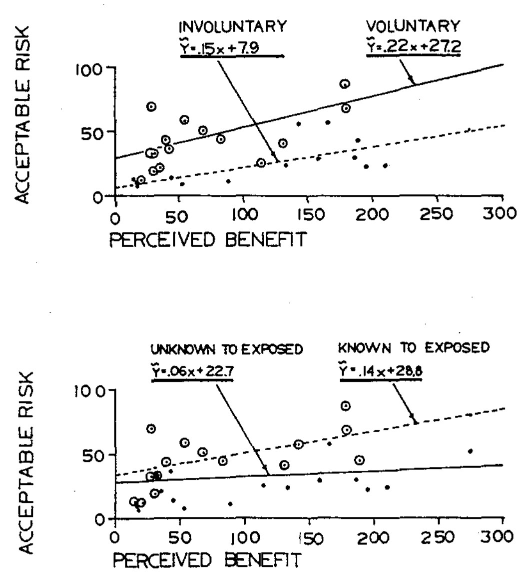

Fig. 6. Relationship between perceived benefit and acceptable risk for voluntary (circled)—involuntary and known (circled)—unknown items.

appropriate if risk levels were made acceptable. This was also reflected in the large negative correlation between voluntariness scores and deviations from the perceived benefit-acceptable risk regression line noted above. Figures like number 6 would show a similar double standard for each of the first six risk characteristics in Table 4.

Factor analysis of risk characteristics

Ratings of the various task characteristics tended to be highly intercorrelated, as shown in Table 5. For example, risks faced voluntarily tended to be known to the exposed individual ($r = 0.83$); new risks tended to be judged less controllable ($r = 0.64$), etc. The intercorrelations were sufficiently high to suggest that they might be explained by a few basic dimensions of risk underlying the nine characteristics. In order to identify such underlying dimensions, we conducted principal components factor analyses (Rummel, 1970) for risk participants and for benefits participants, separately. The unrotated factor loadings for the two groups were so similiar (the mean absolute difference between loadings was 0.05) that they were averaged (Table 6).[1] Two

TABLE 5

Rating Scale Intercorrelations [a]

Scale	Vol.	Immed.	Known to Exposed	Known to Science	Control	New	Chronic	Common	Severity
Voluntariness (1 = voluntary)		0.54*	0.83*	0.75*	—0.76*	—0.65*	0.55*	0.55*	0.06
Immediacy (1 = immediate)			0.78*	0.68*	—0.42	—0.63*	0.16	0.25	—0.22
Known to exposed (1 = known precisely)				0.87*	—0.63*	—0.78*	0.35	0.31	—0.22
Known to science (1 = known precisely)					—0.60*	—0.83*	0.35	0.46	—0.14
Controllability (1 = uncontrollable)						0.64*	—0.63*	—0.64*	—0.24
Newness (1 = new)							—0.46	—0.53*	0.05
Chronic (1 = chronic)								0.60*	0.46
Common (1 = common)									0.63*
Severity of consequences (1 = certain not to be fatal)									

[a] These correlations were computed separately for the risk and benefits group and then averaged (using Fisher's Z transformation).
* $p < 0.001$.

[1] A varimax rotation was applied to these factors, but it produced no improvement in interpretability and will not be discussed.

TABLE 6

Factor Loadings Across Nine Risk Characteristics (Risk and Benefit Subjects Averaged)

Scale	Vol.	Immed.	Exposed	Known to Science	Known to Control	New	Chronic	Common	Severity	λ	Percent of variance accounted for
Factor 1	0.89	0.70	0.88	0.88	—0.83	—0.87	0.62	0.67	0.11	5.30	58.9
Factor 2	0.03	—0.45	—0.39	—0.28	—0.24	0.14	0.55	0.60	0.91	1.90	21.1
Communality	0.79	0.69	0.93	0.86	0.75	0.78	0.69	0.81	0.84		

orthogonal factors appeared sufficient to account for the intercorrelations shown in Table 5. The factor loadings shown in Table 6 indicate the degree to which each risk characteristic correlated with each of the two underlying factors. The first factor correlated highly with all characteristics except severity of consequences. The second factor was associated with severity of consequences and, to a lesser extent, with common/dread and chronic/catastrophic. The communality index in Table 6 reflects the extent to which the two factors accounted for each of the ratings. The communalities were high, indicating that this two-factor solution did a good job of representing the ratings for the nine scales.

Just as each of the 30 items had a (mean) score on each of the nine risk characteristics, we can obtain a score for each item on each factor. These factor scores enable us to plot the 30 items in the space defined by the two factors. As might be expected from the similarity of the factor solutions for risk and for benefit participants, these plots were very similar for the two groups. The mean absolute difference between factor scores was 0.10 for Factor 1 and 0.18 for Factor 2. The correlations between factor scores for the two groups were 0.99 for Factor 1 and 0.98 for Factor 2. Given

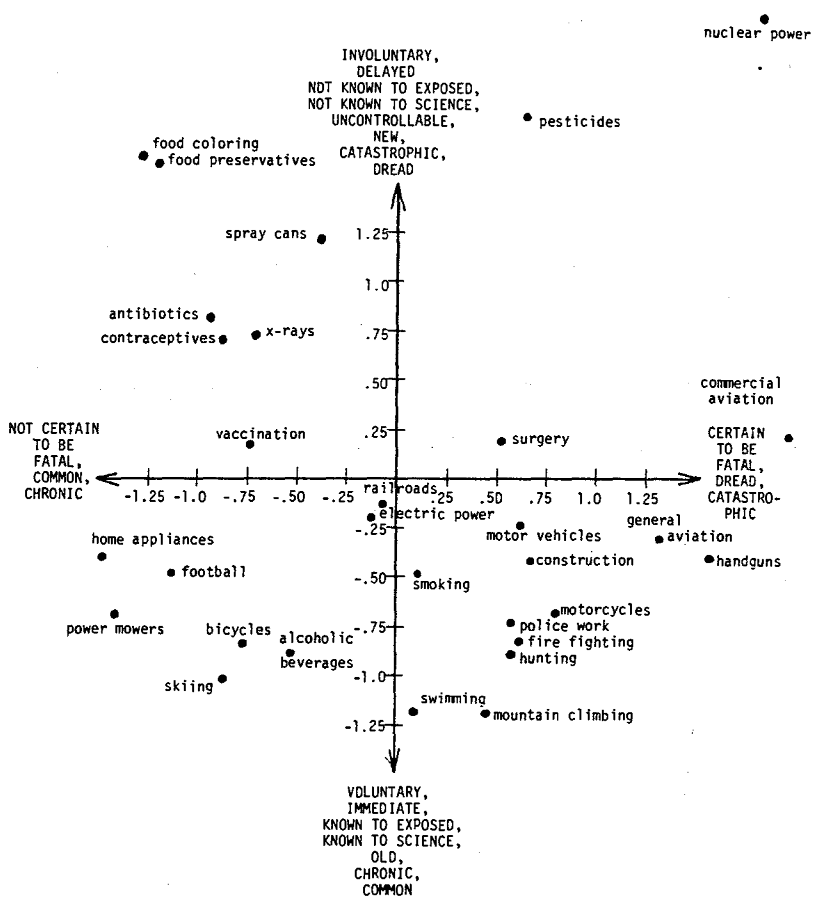

Fig. 7. Location of risk items within the two-factor space.

the extraordinary similarity of the plots for these two independent groups, their factor scores were averaged. Figure 7 plots these scores for the 30 items.

This plot helps clarify the nature of the two factors. The upper extreme of Factor 1 was associated with new, involuntary, highly technological items, which have delayed consequences for masses of people. Items low on the first factor were familiar, voluntary activities with immediate consequences at the individual level. High scores (right-hand side) on Factor 2 were associated with events whose consequences are certain to be fatal (often for large numbers of people) should something go wrong. It seems appropriate to label Factors 1 and 2 as "Technological Risk" and "Severity," respectively.

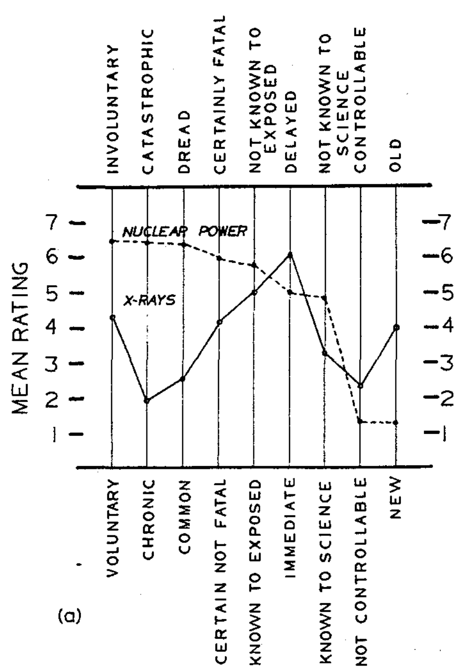

Fig. 8a.

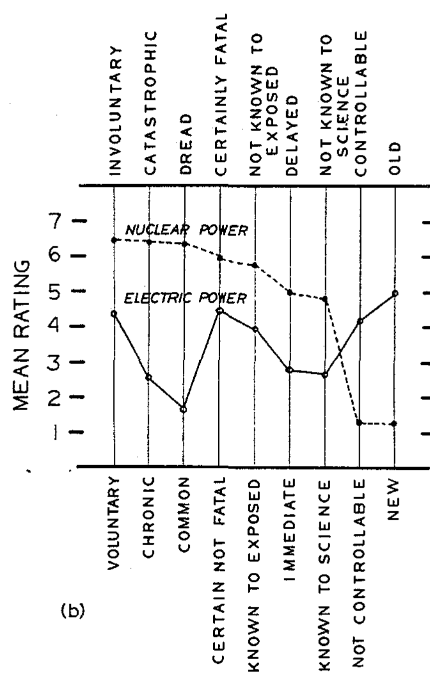

Fig. 8b.

One of the more remarkable features of this factor space was the unique position (isolation) of nuclear power. Clearly, the participants in our study viewed the risks from nuclear power as qualitatively different from those of the other activities. Figures 8a and 8b highlight these differences by comparing the risk ratings for nuclear power and two ostensibly similar technologies, X-rays and non-nuclear electric power. Although X-rays and nuclear power both rely on radioactivity, nuclear power was perceived as markedly more catastrophic and dreaded. For those who believe that nuclear power is just another kind of energy, the discrepancies shown in Fig. 8b should be very surprising. How widely these perceptions are shared by people other than members of the League of Women Voters and their spouses is a matter for future research.

Multivariate Determination of Acceptable Risk Levels

In Fig. 1, Starr's goal was to show a way to predict level of acceptable risk as a function of benefit and voluntariness. Rowe (1977) has done a similar analysis using qualitative aspects of risk other than voluntariness. A generalization of these approaches would be a formula specifying acceptable risk level as a function of benefit and all relevant qualitative aspects of risk. We have done this using the two risk dimensions derived by the factor analysis of the nine qualitative risk scales. A multiple regression equation predicting acceptable risk level as a function of perceived benefit, Factor 1 and Factor 2 yielded a multiple R of 0.76 [$F = 12.2$; df = 3, 26; $p<0.0001$]. This means that we can do a good job of predicting the acceptable risk levels shown in Table 1 from judgments of benefits and several risk characteristics. How such a formula may be used to guide future policy making is also a topic for future research.

A similar analysis was performed on judgments of perceived (current) risk. The multiple R for predicting perceived risk from perceived benefit and the two factor scores was 0.67 [$F = 6.96$; df = 3, 26; $p<0.005$]. However, perceived risk judgments could be predicted just as well using the single qualitative variable "severity of consequences" and ignoring perceived benefit and the other qualitative scales.

Discussion

Methodologically, the main result of this study was that the task we posed to the participants was tractable. That is, it was possible to ask people for complex judgments about difficult societal problems and receive orderly, interpretable responses.

Substantively, the most important findings were:

1. For many activities and technologies, current risk levels were viewed as unacceptably high. These differences between perceived and acceptable risk indicated that the participants in our study were not satisfied with the way that market and other regulatory mechanisms have balanced risks and benefits. Given this perspective, such people may also be unwilling to accept revealed preferences of the type uncovered by Starr as a guide for future action. In particular, the high correlations between perceived levels of existing risk and needed risk adjustment indicated that our participants wanted the risks from different activities to be considerably more equal than they are now. As shown in Table 1, they wanted the most risky item on our list of 30 to be only 10 times as risky as the safest.

2. There appeared to be little systematic relationship between the perceived existing risks and benefits of the 30 activities and technologies considered here. Nor are risks entered into voluntarily perceived as greater than involuntary risks at fixed levels of benefit. Such relationships appeared to emerge in Starr's revealed risk-benefit space.

3. However, there was a consistent, although not overwhelming, relationship between perceived benefit and acceptable level of risk. Despite their desire for more equal risks from different activities, our respondents believed that society should accept somewhat higher levels of risk with more beneficial activities. They also felt that society should tolerate higher risk levels for voluntary, than for involuntary

activities. Thus, they believed that Starr's hypothesized relationships should be obtained in a society in which risk levels are adequately regulated. In addition, other characteristics of risk besides voluntariness, namely perceived control, familiarity, knowledge, and immediacy, also induced double standards for acceptable risk. Thus, these expressed preferences indicate that determining acceptable risk may require consideration of other characteristics besides benefits.

4. The nine characteristics hypothesized by various authors to influence judgments of perceived and acceptable risk were highly intercorrelated. They could be effectively reduced to two dimensions. One dimension apparently discriminated between high- and low-technology activities, with the high end being characterized by new, involuntary, poorly known activities, often with delayed consequences. The second dimension primarily reflected the certainty of death given that adversity occurs. Consideration of these two factors in addition to perceived benefit made acceptable risk judgments highly predictable. Conceivably, policy makers might use such relationships to predict public acceptance of the risk levels associated with proposed technologies.

Given the contrasts between our study and Starr's, the question arises, "Who is right?" We believe that neither approach is, in itself, definitive. The particular relationships that Starr uncovered were based upon numerous ad hoc assumptions and applied to only a small set of possible technologies. Our own study used but one of the psychophysical measurement procedures possible, applied to a rather special participant population.[2] We are, at present, engaged in additional studies employing different types of respondents and different judgment methods. Answering the question "How safe is safe enough?" is going to require a multi-method, multidisciplinary approach, in which the present work and Starr's are but two components.

Balancing the results of these various approaches also depends upon one's conceptualization of the policy-making process. A definitive revealed-preference study would be an adequate guide to action only if one believed that rational decision making is best performed by experts formalizing past policies as prescriptions for future action. A definitive expressed-preference study would be an adequate guide only if one believed that people's present opinions should be society's final arbiter and that people act on their expressed preferences. The obvious reservation that many people would have about the former approach is that it is highly conservative, enshrining current economic and social relationships; an obvious problem with the latter approach is that it allows people to change planning guidelines at will, possibly resulting in social chaos.

For most people, presumably, both present opinions and past behavior are relevant to social policy. The believer in expressed preferences cannot ignore existing economic arrangements. On the other hand, the public will resist even the best-laid

[2] Preliminary results indicate that the risk-benefit relationships obtained with the League of Women Voters subjects replicated almost exactly when a group of university students made the same sorts of judgments. A second study asked students to judge acceptable risks directly, instead of using an adjustment factor. The direct ratings correlated about 0.77 with ratings produced by another group using the indirect adjustment method of the present study. These results indicate an encouraging degree of cross-method consistency.

plans if they feel that policy makers have not adequately considered their desires. Assume that future research finds that a representative sample of properly informed citizens, queried by means of appropriate methods, evaluates the seriousness of hazards not only by their statistical and "economic" risks, but also according to qualitative features like voluntariness and controllability. The legitimacy of these desires will have to be explored and debated. For example, implementing a double standard for voluntary risks may prove, upon analysis, to be acceptable while the desire to make dreaded technologies especially safe may be found to have unreasonable consequences. Even if the public's desires are ignored, either with or without analysis, there is no guarantee that they will go away. Pressure on politicians and regulators may force laws based on more "rational" economic considerations to be implemented in accordance with these "irrational" desires. Indeed, the current functioning of our regulatory system might be better understood as a partial reflection of such pressures.

Although we have deemphasized the substance of our respondents' judgments about specific technologies in order to concentrate on more general relationships between those judgments, such opinions from members of the League of Women Voters are quite likely to appear in regulatory hearings and elsewhere. If League members believe that nuclear power has low benefit relative to its level of risk, it is as much a political fact of life as the League members' failure to see any systematic tradeoff between existing risks and benefits.[3]

The present study raised several questions worthy of further investigation. One intriguing finding was that people viewed current risk levels as more acceptable after they had ordered current benefits in depth (Table 1, columns 3 and 4). Does this imply that the way technologies are presented, say, in regulatory hearings, can affect the way in which they are evaluated? More research is needed into how to present the public with the information needed to give new technologies a fair hearing.

A second question is triggered by the observed inverse relationship between perceived risks and benefits. Could this have occurred because participants in the benefit group were unable to estimate gross benefits rather than net benefits? If people in the benefit group did take risk into consideration, high-risk activities would have been rated as relatively lower in benefit and low-risk activities would have been viewed as relatively higher in benefit, much like the observed pattern. Future work should consider the advantages of having people judge multiple aspects of benefit (e.g., economic aspects, physical and mental health, convenience, etc.) separately. These could then be weighted and amalgamated into an overall, multi-attribute measure. This approach may reduce or eliminate possible contamination from the risk side.

Finally, what is the relationship between these attitudes about risk and people's responses to measures designed to ameliorate risks? If people believe that motor

[3] In November of 1976, half a year after distribution of our questionnaire, Oregon voters decided the fate of a nuclear safeguards ballot measure that if passed would have curtailed, and perhaps stopped, the development of nuclear power in Oregon. In response to a survey preserving their anonymity, 95% of the participants in our study indicated voting in favor of the safeguards measure (i.e., against nuclear power) compared with 42% supporting it statewide. Thus, the voting behavior of our League subjects matches the anti-nuclear sentiments they expressed in their risk and benefit judgments.

vehicles should be five times safer, does this mean that they would accept any immediate, Draconian step designed to attain that goal? Does it mean that a fivefold reduction in risk is a long-term goal for society and that meaningful (but not necessarily drastic) steps should be taken until that goal is reached, or does it mean that the adjustment ratios expressed here only measure relative concerns about the risk levels of various activities? A more behaviorally relevant scale of acceptability should be developed, with clearer implications for regulatory actions.

REFERENCES

Acton, J. P. (1973). "Evaluating public programs to save lives: The case of heart attacks," *Rand Corporation Report R–950–RC*, January.

Golant, S. and Burton, I. (1969). "Avoidance response to the risk environment," Natural Hazards Research Working Paper No. 6, Dept. of Geography, University of Toronto.

Green, C. H. (1974). "Measures for safety." Unpublished manuscript. Center for Advanced Study, University of Illinois, Urbana.

Fischhoff, B. (1977). "Cost-benefit analysis and the art of motorcycle maintenance," *Policy Sciences*, **8**, 177–202.

Kates, R. W. (1975). "Risk assessment of environmental hazard," *SCOPE Report 8*, International Council of Scientific Unions, Paris, France.

Lichtenstein, S., Slovic, P., Fischhoff, B., Combs, B. and Layman, M. (1978). "Perceived frequency of low-probability lethal events," *Journal of Experimental Psychology: Human Learning and Memory*, in press.

Linnerooth, J. (1975). "The evaluation of life saving: A survey," *Research Report 75–21*, International Institute for Applied Systems Analysis, Laxenburg, Austria, July.

Liska, A. E. (Ed.), (1975). *The Consistency Controversy*. New York: Wiley.

Lowrance, W. W. (1976). *Of Acceptable Risk*. Los Altos, Calif.: Wm. Kaufman, Inc.

Maynard, W. S., Nealey, S. M., Hébert, J. A. and Lindell, M. K. (1976). "Public values associated with nuclear waste disposal," *Report BNWL–1997 (UC–70)*, Battelle Memorial Institute, Human Affairs Research Center, Seattle, Washington, June.

Otway, H. (1975). "Risk assessment and societal choices," *Research Memorandum 75–2*, International Institute for Applied Systems Analysis, Laxenburg, Austria, February.

Otway, H. J. and Cohen, J. J. (1975). "Revealed preferences: Comments on the Starr benefit-risk relationships," *Research Memorandum 75–5*, International Institute for Applied Systems Analysis, Laxenburg, Austria, March.

Otway, H. J., Maderthaner, R. and Guttman, G. (1975). "Avoidance response to the risk environment: A cross cultural comparison." *Research Report 75–14*. International Institute for Applied Systems Analysis, Laxenburg, Austria, May.

Otway, H. J. and Pahner, P. D. (1976). "Risk assessment," *Futures*, **8**, 122–134.

Rappaport, E. (1974). "Economic analysis of life-and-death decision making." Appendix 2 in *Report No. Eng 7478*, School of Engineering and Applied Science, UCLA, Nov.

Rowe, W. D. (1977). *An Anatomy of Risk*. New York: Wiley.

Rummel, R. J. (1970). *Applied Factor Analysis*. Evanston: Northwestern Univ. Press.

Siegel, S. (1956). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

Slovic, P., Fischhoff, B. and Lichtenstein, S. (1977). "Behavioral decision theory," *Annual Review of Psychology*, **28**, 1–39.

Starr, C. (1969). "Social benefit versus technological risk," *Science*, **165**, 1232–1238.

Starr, C. (1972). "Benefit-cost studies in sociotechnical systems." In Committee on Public Engineering Policy, *Perspective on Benefit-Risk Decision Making*. Washington, D.C.: National Academy of Engineering.

Starr, C., Rudman, R. and Whipple, C. (1976). "Philosophical basis for risk analysis," *Annual Review of Energy,* **1**, 629–662.

Torrance, G. (1970). "Generalized cost-effectiveness model for the evaluation of health programs," *McMaster University Faculty of Business Research Series, No. 101.*

Wyler, A. R., Masuda, M. and Holmes, T. H. (1968). "Seriousness of illness rating scale," *Journal of Psychosomatic Research,* **11**, 363–374.