

## How Serious Are Expressions of Protected Values?

Jonathan Baron and Sarah Leshner  
University of Pennsylvania

People think that some things that they value should be protected from trade-offs with other things. For example, people think that no economic gain is great enough to justify clear-cutting old-growth forest. The authors probed the stability of these protected values (PVs) in several ways. Subjects were asked to think of counterexamples, and this had some effect on PVs. Subjects were then asked how they would resolve conflicts between 2 PVs. Resolutions did not differ from those between other values, but subjects tended to feel that conflicts between PVs did not occur in reality. Despite people's claims that PVs are unchanged by variation in quantity, expression of PVs was reduced when the magnitude or probability of the violation of a PV was smaller. In summary, PVs appear to be labile and amenable to challenge. Despite earlier concerns, apparent PVs may not always preclude the use of valuation measures in cost-effectiveness analysis or negotiated agreement on controversial issues.

People think that some of their values are protected from trade-offs with other values (Baron & Spranca, 1997; Tetlock, Lerner, & Peterson, 1996). Many of these values concern natural resources, such as species and pristine ecosystems. People with protected values (PVs) for these things do not think they should be sacrificed for any compensating benefit, no matter how small the sacrifice or how large the benefit. The term *value* here is used to mean utility, that is, the measure of desirability that decisions are meant to increase (Baron, 2000). In an economic sense, when values are protected, the marginal rate at which one good can substitute for another is infinite. For example, no amount of money can substitute for a type of environmental decline.

Several researchers have noted that PVs cause problems for quantitative elicitation of values, as is done in cost-benefit analysis or decision analysis (Baron, 1997, 1999; Bazerman, Moore, & Gillespie, 1999). Methods used in such elicitations include contingent valuation, decision analysis, and conjoint analysis (Baron, 1997; Calfee & Winston, 1998; Keeney & Raiffa, 1993; Louviere, 1988; Mitchell & Carson, 1989). In principle, such methods permit optimal allocation of scarce resources (Gregory, Lichtenstein, & Slovic, 1993; Keeney & Raiffa, 1993; Portney, 1994).

However, application of these methods often requires measurement of values—in terms of money or utility. When a value has infinite utility for some person, it swamps everything else and makes this sort of analysis impractical. In practical terms, people cannot spend all their resources on protecting the environment, saving human lives, protecting human rights, or any one thing. People must make trade-offs. Suppose the measure of value is willingness to pay. If one tries to find the average willingness to

pay more taxes, say, to save a forest and if some people say that the forest has infinite value, the average will be infinite, regardless of what others say. Increased expenditures on one good involves sacrifices of other goods, such as human life and health, and other people may claim absolute values for these. If that happens, no decision is possible on the basis of measures of values. Even if only one side of the equation has an absolute value, it could almost never be honored in practice. Thus, the two problems are as follows: (a) the possibility that one person could dominate a decision by expressing an absolute value and (b) the possibility that people with conflicting absolute values can make a decision impossible. The problems can happen with any measure of value that allows expression of PVs.

Some writers see the problem of refusal to make trade-offs as the basis of a philosophical objection to cost-benefit analysis. Anderson (1993), for example, criticized cost-benefit analysis because it

ignores the possibility that goods such as endangered species may be specially valued as unique and irreplaceable higher goods. The distinction between higher and lower goods, which supports norms that prohibit certain tradeoffs between them, plays no part in the analysis. (pp. 193–194)

Other writers speak of “protest responses” to questions about economic values of resources, such as the dollar value of environmental preservation (Mitchell & Carson, 1989, pp. 32–34). When respondents refuse to answer questions about the dollar value of a natural resource such as a forest, this is seen as a methodological problem in survey design, to be overcome by such techniques as asking people yes-no questions about whether or not they would accept a certain amount of money.

PVs also cause problems for negotiation. Governments often try to settle environmental (and other) disputes through negotiation among interested parties or their representatives (Thompson & Gonzalez, 1997). When parties to the disputes have PVs, negotiation gets stuck. The task of a mediator is to look for possible compromises, and some of these involve the sacrifice of PVs. In this case, of course, false expressions of PVs are often used as negotiating ploys, but sometimes these expressions may stem from

---

Jonathan Baron and Sarah Leshner, Department of Psychology, University of Pennsylvania.

This research was supported by National Science Foundation Grant SBR95-20288.

Correspondence concerning this article should be addressed to Jonathan Baron, Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia, Pennsylvania 19104-6196. Electronic mail may be sent to baron@psych.upenn.edu.

the same sources as they do in responding to surveys about values. These seem not to be a matter of mere posturing (Baron & Spranca, 1997, Experiment 1).

At issue in this article is whether PVs are as absolute as they seem to be when people first express them, and, if not, what kinds of challenges and probes can lead people to express willingness to make a trade-off with other values (typically economic). Our concern is not behavior. Surely, people who endorse PVs violate them in their behavior, but these violations do not imply that the values are irrelevant for social policy. People may want public decisions to be based on the values they hold on reflection, regardless of their behavior. When people learn that they have violated some value they hold, they may regret their action rather than revise the value.

In saying this, we express a concept of values (or utilities) as criteria for evaluating states of affairs. Values are reflectively endorsed. They are the result of thought and are, in this sense, constructed, much in the way that concepts are the result of reflection. Values are not simply desires, and very young children may properly be said to have no values at all, in the sense at issue. Values are important because they are people's best judgments of the goodness of outcomes. Insofar as governments or other organizations seek to produce good outcomes, the ultimate judgments of those outcomes are values.

If values are seen as constructed (like concepts), we can ask whether they are constructed well, just as we can ask whether concepts are formed well. It is possible that PVs result from the same kind of unreflectiveness that leads to incorrect or overgeneralized concepts (Baron, 1973). People may agree with the claim that "all apples are red" without pausing to consider counterexamples. Likewise, they may endorse the statement that "no benefit is worth the sacrifice of a pristine rainforest" without thinking much about possible benefits (e.g., a cure for cancer or malaria). Or, when people say that they would never trade off life for money, they may fail to think of extreme cases, such as crossing the street (hence risking loss of life from being hit by a car) to pick up a large check or failing to increase the health care budget enough to vaccinate every child or screen everyone for colon cancer. Such unreflective overgeneralizations provide one possible avenue for challenging PVs to make compromise and trade-offs possible. If PVs are unreflective in this way, then PVs should yield to simple challenges.

This hypothesis is outwardly similar to one that has been tested in social-psychological studies of values. Such studies have found that values were labile and easily changed by asking people to think about their values (e.g., Maio & Olson, 1998) or by asking about the application of the values to a particular context (e.g., Seligman & Katz, 1996). The difference between the present approach and these social-psychological studies is in the measure of values. The previous studies used surveys based on that of Rokeach (1973), which asked subjects to rate principles such as environmental protection or hedonism as "guiding principles in your life" on a 9-point scale from *opposed to my values* or *not important* to *extremely important*. Answers to such questions have no clear implications about trade-offs, and they are difficult to interpret as statements about rates of substitution between one value and another (e.g., how much self-gratification one would sacrifice for a given amount of environmental protection). The lability of these ratings may reflect how items are interpreted

rather than how they would be traded off. By contrast, questions about PVs of the form used here are explicitly about trade-offs. Specifically, they concern whether the subject thinks of a rate of substitution as infinite or not.

Experiments 1–4 address the possibility that PVs are unreflective overgeneralizations. Subjects answered questions about whether they would regard certain outcomes, such as "electing a politician who has made racist comments," as so much against their values that no benefit would be sufficient to justify actions that caused such outcomes. Then, when values were protected in this way, we challenged them by asking the subjects to think of counterexamples. PVs sometimes change in response to such challenges. Experiment 4 also found that the effects of counterexamples can transfer to measures of omission bias (Ritov & Baron, 1990), the bias toward harm caused by omissions when that is pitted against harm caused by acts.

Another possibility is that people have not thought much about what happens when PVs conflict with each other. In Experiment 5, we ask people specifically about such conflicts, which people seem to find unproblematic. They are thus willing to trade off PVs when they conflict with other PVs.

Experiments 6 and 7 explore what happens when harm (that goes against a PV) is probabilistic and when it varies in amount. If (as we find), PVs are not honored when the probability and magnitude of harm is low enough, then this suggests a way in which we can measure trade-offs. It may also suggest that PVs are, in a sense, unreflective. When people say that a value is absolute, they seem to have in mind a violation of a certain magnitude and probability.

In studying PVs, we rely somewhat on previous findings that suggest that PVs concern rules about action, irrespective of their consequences, rather than consequences themselves. What counts as a type of action (e.g., lying) may be defined partly in terms of its consequence (false belief) or intended consequence, but the badness of the action is not just that of its consequences, therefore it has value of its own.

First, people think that their PVs should be honored even when their violation has no consequence at all. People who have PVs for forests, for example, say that they should not buy stock in a company that destroys forests, even if their purchase would not affect the share price and would not affect anyone else's behavior with respect to the forests. This is an *agent-relative* obligation, a rule for the person holding the value that applies to his own choices but not (as much) to his obligations with respect to others' choices. Thus, it is better for him not to buy the stock, even if his not buying it means that someone else buys it.

Second, PVs are insensitive to quantity. People who hold a PV for forests tend to say that it is just as bad to destroy a large forest as a small one (Lim & Baron, 1997; Ritov & Baron, 1999). They say this more often than they say the same thing for violations of nonprotected values (NPVs).

Third, PVs tend to apply to acts, not omissions, and to trade-offs with gains in other values, not losses (Ritov & Baron, 1999). People with PVs show more omission bias: They are less willing than others to destroy one forest to prevent destruction of other forests, holding the destruction constant. It is not the destruction that matters to them so much as the act of destroying. PVs are also somewhat limited to trade-offs with gains. Some people are unwilling to destroy a forest to promote "economic" gain, but they

are willing to destroy it to prevent a loss. Of course, what counts as a gain or loss is relative and is easily manipulated (e.g., Shafir, Diamond, & Tversky, 1997). Such differences in description are relevant to classifications of acts, even though they are not relevant to evaluation of consequences. In summary, PVs seem to be deontological.

### Experiment 1

The hypothesis for this experiment was that claims to have PVs are sometimes the result of incomplete thinking. For example, a person who says that an old-growth rainforest should not be cut for anything may not have tried very hard to think of benefits that may be great enough to justify such an action. When people are forced to think about such benefits, they may backtrack on their original claims.

Some previous findings suggest that people endorse general and absolute statements about values and then later endorse particular statements that conflict with the general ones. In a survey of U.S. attitudes toward civil liberties, McClosky and Brill (1983) found apparent conflict between general and particular statements. For example, 89% of the respondents agreed with the statement "I believe in free speech for all no matter what their views might be" (p. 50), but only 29% agreed that "Groups like the Nazis and the Ku Klux Klan [should] be allowed to appear on public television to state their views" (p. 70).

It is not clear, however, whether the 60% difference represents truly conflicting answers. People may think that free speech does not imply the right to speak on public television, or in any particular place, as long as some avenue of expression is available. Here, we ask subjects to provide their own counterexamples, thus ensuring that they perceive the examples as conflicting.

Another difference between the present study and that of McClosky and Brill (1983) is that the general statements we present are designed to assess absoluteness in the willingness to trade off one value with any other value. Thus, instead of saying "no matter what their views . . .," our questions are of the form "no matter how great the benefits . . ." without specifying what the benefit may be. Our major concern is the willingness to make trade-offs.

### Method

Thirty-five students from the University of Pennsylvania (Penn) and 15 students from the Philadelphia College of Pharmacy and Science (PCPS; now renamed University of the Sciences in Philadelphia) served as subjects. (Although 25 of the Penn students were known to Sarah Leshner, their results did not differ from the other Penn students.) Subjects were recruited through posters and electronic news groups. Each person was paid \$3.00 for participating in a 20-min interview.

Thirteen situations were presented (by Sarah Leshner) to identify which values subjects believed were protected. The questionnaire explained that PVs are those that people hold so dear that they think the consequences of acting against them are always worse than any benefits they may gain. A list of 13 situations in which some people may not think there could ever be a trade-off great enough to allow it to happen was read to subjects. For each one, subjects chose "A" if they thought the situation was never acceptable, no matter how great the benefits, and "B" if it would be acceptable if it led to some sort of benefits that are great enough. The items were as follows: buying clothes from a manufacturer that uses young children at sewing machines, electing a politician who has made racist comments, eating tuna if tuna fishing boats kill many dolphins every year,

building a dam that threatens many fish with extinction, cutting all the trees in an old-growth forest, raising children's IQ through genetic engineering, sterilizing retarded women, exposing workers at a steel mill to hazardous chemicals, giving money to an organization where the president was caught embezzling, using baby rabbits to test cosmetic products, a manufacturing company dumping its trash in a local river, building a new public water system with pipes found to contain lead, and doctors removing dying patients' organs without their consent.

After all 13 situations were presented, the interviewer used two strategies to induce trade-offs where subjects had said they would never be permissible. The first tactic was to pit PVs against each other by asking what would happen if the subject's PVs came into conflict. For example, if a subject reported an absolute prohibition against buying clothes from a manufacturer that uses child labor, the interviewer would ask what would happen if this conflicted with some other value on the list that the subject also said was absolute. The second tactic was simply to ask the subject to imagine counterexamples: "It is possible that some of these answers were not really protected values. Please go back and try to think of some benefits that would make each of these actions acceptable." We then invited the subject to change his or her initial answers. Subjects were also probed for additional PVs to be used as hypothetical situations in the next study.

### Results

Table 1 shows the percent of PVs before the request to think of counterexamples, across all subjects, by item. The second column shows the percent of those who had PVs but could think of counterexamples. The average proportion of counterexamples that led to changes was 23%, but this was only 10% of all the PV responses (because many did not yield counterexamples—these proportions were computed for each subject and then averaged across subjects).

The PCPS students generally thought of fewer counterexamples than the Penn students,  $t(48) = 3.17, p = .0027$ , two-tailed (all tests are one-tailed unless noted), but otherwise the groups did not differ significantly. (In all studies reported, no other effects of age, sex, or student status were significant unless noted.)

The attempt to challenge PVs by asking how subjects would resolve conflict yielded little information. The two most frequent recommendations were human life must always prevail and choose

Table 1  
Percent Protected Values (PVs) in Experiment 1 and  
Percent of PV Responses for Which Subjects  
Could Think of Counterexamples

Question	PV before challenge	Counterexamples
Child labor	36	71
Racist politician	52	58
Tuna	28	79
Dam building	36	72
Tree cutting	58	69
Raising children's IQ	50	52
Sterilize retarded women	66	48
Steel mill workers	80	46
Nonprofit organization	50	54
Cosmetics on bunnies	44	41
Dump trash in river	90	40
Lead pipes	88	39
Doctors removing organs	64	59
Total	57	53

the one that does the least damage. Experiment 3 investigates conflict more systematically.

### Experiment 2

Experiment 2 asked for general statements about the existence of counterexamples rather than for specific counterexamples. We gave subjects several options for indicating what kinds of counterexample might exist. The purpose of this experiment was to find out whether PVs still remained when all options were available in this way.

We based these options on answers to Experiment 1 and also arguments made in the philosophical literature (e.g., Hare, 1981). In particular, utilitarian philosophers since Mill have argued that even though utilities are truly commensurable, people should sometimes follow rules and make no exceptions. These cases occur when the exceptions are so difficult to recognize correctly when they happen (in part because they are rare) that, in any given case, people would do better to follow the rule. For example, most political terrorists see themselves as acting on behalf of the greater good, but, in hindsight, essentially none have been correct, therefore it is better never to commit terrorism even when it seems to you that the trade-off is worth making.

We also allowed subjects to say that something was unacceptable as a rule but had exceptions, and we gave other specific examples of cases in which violation of a value might be acceptable. Another response option said that exceptions may be imagined but do not occur in the real world. This would not be a true PV, but people may appear to have PVs when they believed this. People with such beliefs would, in principle, yield to presentation of an actual counterexample in the real world.

This questionnaire, and others, were completed on the World Wide Web. Subjects found out about the study from links in a variety of web pages, including one advertising "free stuff on the internet." They had to provide an address and social security number to be paid. The questionnaires included a variety of checks to make sure that responses were serious (although there were more of these in subsequent experiments). Several articles in Birnbaum (2000) compared World Wide Web questionnaires with paper questionnaires given with the experimenter present. These articles found either no difference in the responses or (in a few cases) evidence of more careful responding on the World Wide Web. Studies on the World Wide Web also yield a greater variety of subjects and allow more sensitive tests for the effects of such variables as age and student versus nonstudent status. (This is more true in studies done later, when more nonstudents were exploring the World Wide Web.)

### Method

There were 50 subjects (40% female, 86% students, median age = 21 years, range = 17–49), recruited through internet news groups and paid \$5.00 for completing a questionnaire on the World Wide Web. The same situations were used as in Experiment 1, in the same fixed order, with the omission of "eating tuna if tuna fishing boats kill many dolphins every year" and the addition of "eating a dying companion," "dropping an atom bomb," and "executing minors." These questions were obtained from subjects' answers in Experiment 1.

The instructions asked subjects to indicate which code numbers applied to each value. (They could endorse more than one.) The codes, with the

percent of cases on which each code was endorsed in parentheses, are as follows. This is acceptable if

- A1. There are economic benefits. (11)
- A2. It saves a human life. (30)
- A3. It improves living conditions. (17)
- A4. It prevents a global problem. (23)
- A5. It is for a good moral cause. (14)
- A6. It prevents more of the same thing. (20)

This is unacceptable because

- U1. I cannot imagine any situations in which this is acceptable. (38)
- U2. I can imagine situations in which the benefits are great enough to justify this, but these situations do not happen in the real world. (7)
- U3. There are situations in the real world in which the benefits are great enough, but people can not recognize these situations, so it is best never to do this. (9)
- U4. This is unacceptable as a general rule, but we should make exceptions to it if we are sure enough. (28)

### Results

The percent of times each response was endorsed is shown above. Despite the opportunity to express qualifications, 38% of the responses indicated true PVs (U1). Even when we omit cases in which the subject endorsed A1–A6 or U4 in addition to U1 (thus admitting some exceptions, possibly through carelessness), the number of U1 responses was 34%.

Subjects endorsed U2 or U3 in only 13.5% of the cases (endorsing both in 1.9%). These do not appear to be a major source of potential counterexamples. U4, however, was endorsed 28% of the time (26% if we remove cases in which U1 was also endorsed). People think of rules as permitting exceptions, but these were not the same rules that led subjects to endorse proper PVs (U1) because only 2.5% of responses endorsed both of these. In summary, although subjects do recognize the existence of imaginary and unrecognizable counterexamples (U2 and U3), these do not appear to be of major value as a way to encourage people to admit trade-offs.

### Experiment 3

The results of Experiment 1 suggest that PVs are occasionally unreflective overgeneralizations. PVs sometimes yielded to counterexamples. Experiment 3 combined the methods of Experiments 1 and 2, using both challenges and a wide range of response options. This allowed us to ask whether counterexamples are found even with the wider range of options used in Experiment 2.

### Method

Forty-six subjects completed a World Wide Web questionnaire for \$5.00. Two were not challenged on any item. Of the remaining 44, 41% were male, 86% were students, and the median age was 20 years (range = 13–50).

We used the 15 situations from Experiment 2 and the same instructions. Subjects could mark multiple reasons in each category but could not mark U1 in combination with A1–A6. The order of presenting the items was fixed.

The questionnaire (written in JavaScript) presented a middle section in which items with no "acceptable" (A1–A6) response were challenged. For these situations, the subject was told to reconsider and "try to think of a specific case where the action would be acceptable in the real world. If you can think of one, write in a brief description. If not, write 'none.'" For

example, a subject might have chosen U1 ("I cannot imagine any situations in which the benefits are great enough to justify this") for the question about building a water system with lead pipes. The JavaScript program then encouraged the subject to try to imagine a reason this could be acceptable, such as if there was no better alternative. (Because of a programming error, the first 13 subjects were challenged on some items that were not initially protected. We ignored these responses.) Subjects were then presented with the 15 situations again and told to "try to take into account the examples you have just thought of, if any." They could change any of their original answers.

## Results

Of the values protected on the first trial and challenged, 34% yielded counterexamples. Of the challenged items that yielded counterexamples, 40% changed to being unprotected on the second trial, whereas 5% of those that did not yield counterexamples became unprotected,  $t(24) = 5.04$ ,  $p < .0001$ , for the subjects who had data in both categories. Hence, the challenges were effective when subjects were able to think of counterexamples.

On the first trial, 57% of the values were protected (defined as no "acceptable" [A1–A6] responses). On the second trial, after the challenges, 54% were protected, and the overall change was not significant. The total reduction in PVs caused by the challenges was mitigated by changes in the opposite direction for the 43% of items that were not challenged.

It is important to note that the proportion of PVs was the same as in Experiment 1 (57%) but that the proportion of counterexamples was lower (34% vs. 53%). This suggests that the extra response options eliminated some PVs that would otherwise have yielded counterexamples. However, the subject populations were different too.

## Experiment 4

Experiment 4 asked subjects to try to think of counterexamples to their expressions of PVs, and it examined the effect of this thinking on the PVs that they thought about, as well as PVs that they did not think about. It also asked whether omission bias is reduced for the PVs thought about and in general. Ritov and Baron (1999) found that omission bias was higher when PVs were involved. Omission bias was measured by asking hypothetical questions of the form, "Would you take an action that would cause the extinction of 2 endangered species in order to prevent the extinction of 20 species?" and, more generally, "What is the largest number of extinctions that you would accept in order to prevent 20?" Typically, the threshold number is far less than 19. Ritov and Baron found that it was lower for subjects who had PVs for species than for those who did not. This implies that PVs apply more to actions than omissions. People with PVs were willing to tolerate more harm from their omissions than from their actions. If PVs are reduced by thinking of counterexamples, we may expect the threshold to increase as well. We interpret an increase in the threshold as a reduction in omission bias.

We tested four groups in an experiment on the World Wide Web. First, a control group simply answered questions about PVs and omission bias twice for eight items. One experimental group was asked to think of counterexamples to all eight items. Another group was asked to think of counterexamples to the first four, and

another to the second four. The experimental groups then answered all the questions again.

## Method

The 63 subjects included 38 females and 11 nonstudents. Their ages ranged from 18 to 68 years ( $Mdn = 21$ ). Subjects were assigned to conditions by their birthdays, with approximately equal numbers in the four conditions.

Omission bias was measured with a threshold question such as the following:

1. Vaccination. A vaccination program will prevent 1,000 children from dying from an epidemic of a new infectious disease. The vaccine itself will kill 100 children because it sometimes causes the same disease.

Would you initiate the program? [Checkboxes for no and yes.]

What is the largest number of children killed by the vaccine at which you would initiate the program?

Items were omitted from analysis when the yes–no question disagreed with the numerical answer. (Two subjects' data for all threshold questions were omitted because more than half of the items had such disagreements. For the remaining subjects, 8% of the items were omitted this way.) The amount of omission bias was defined by that answer, with greater bias when the answer was lower. (Almost all were less than 1,000.)

The other items were, in brief form, as follows:

2. River diversion. As a result of a dam on a river, 20 species of fish are threatened with extinction. By opening the dam for a month each year, you can save these species, but 2 species downstream will become extinct because of the changing water level.

3. Starvation. A convoy of food trucks is on its way to a refugee camp during a famine in Africa. (Airplanes cannot be used.) You find that a second camp has even more refugees. If you tell the convoy to go to the second camp instead of the first, you will save 1,000 people from death, but 100 people in the first camp will die as a result.

4. Cutting forests. A logging company has the rights to 1,000 square miles of old-growth forest. The company is willing to trade this land for 100 square miles of similar land, now part of a national park. You can give the smaller area to the company and make the larger area into a national park. . . . The logging company will cut all the trees in whichever area it owns.

5. Civil rights. A prosecutor who works for you has just begun to work on a case involving discrimination against blacks in applications for mortgages. 10 instances of discrimination are involved. A similar case comes to your attention, and you have no other prosecutor to assign. The new case involves 100 instances of discrimination.

6. Medical assistance. Funds for cancer treatments for poor people are limited. Right now, the program you run covers an expensive treatment that cures 100 people each year. If you stop covering this treatment completely, you can use the same money for another, less expensive, treatment that can cure 1,000 people of an equally bad kind of cancer.

7. Child labor. You are a buyer for a large clothing chain. You are buying clothes from supplier A in a poor country, who employs 1,000 ten-year-old children to work at sewing machines. You can switch to supplier B. You know that B will start factories that will employ 100 ten-year-olds who are now not working.

8. Dolphins. You are in charge of regulating some tuna fishing boats. The boats now fish in an area in which they catch 1,000 dolphins per year. The dolphins die in the nets, unable to breathe. You could

require the boats to move to another area, where they would catch just as many tuna, but 100 dolphins instead of 1,000.

The PV question asked subjects to check one of the two responses in items, such as causing the extinction of fish species: This is acceptable if it leads to some sort of benefits (money or something else) that are great enough. This is not acceptable no matter how great the benefits. The counterexample instructions, given after all the PV responses, were as follows:

For many items in the last part, many people say "This is not acceptable no matter how great the benefits." Possibly, in some of these cases, people could imagine possible benefits great enough to justify the action—as a matter of policy—if they tried to do so.

For each of the following cases, please try to imagine and write down (briefly) some benefit that would make this item acceptable. If you can imagine several answers, write the one that is most realistic. If you simply cannot imagine any answer, say so. Please do this for all the items listed, regardless of your previous answers.

Space was provided for a written answer. One group was given the stems for the PV questions for all eight items, one group was given the first four, one group was given the second four, and one group was not given this part. All subjects then answered the original questions again—PV questions first, then omission-bias questions. They were not told why they did this.

## Results

The correlation between PVs and omission bias (found by Ritov & Baron, 1999) was replicated. Omission bias was measured as the response to the threshold question, with lower responses indicating more bias. The mean correlation between threshold and PVs (for the first response to each) was  $-.18$ ,  $t(7) = 4.77$ ,  $p = .0010$ , across the eight items. In addition, the correlation across subjects between mean threshold and proportion of PVs was  $-.35$  ( $p = .0023$ ). It is possible that these correlations could be increased by the effect of answering one question on answering another one.

Table 2 shows mean proportions of PVs and threshold proportions before and after the challenge as a function of the experimental condition. The challenge—thinking of counterexamples—reduced PVs and also increased the threshold for omission, although the latter result was not found in all analyses. The effect on PVs appeared in several analyses. First, the two groups in which half of the items were challenged allow a within-subject

**Table 2**  
*Mean Proportion of Protected Values (PVs) and Mean Threshold Proportions for Omission, Before and After Challenge, in Experiment 4*

Challenge	Challenged items		Unchallenged items	
	Before	After	Before	After
PVs				
No challenge			.27	.28
Half challenged	.47	.34	.46	.47
All challenged	.55	.37		
Thresholds				
No challenge			.51	.52
Half challenged	.41	.42	.40	.36
All challenged	.46	.51		

comparison. The proportion of PVs in the challenged items declined from 48% to 34%, from the first test to the second. The proportion of PVs in the unchallenged items remained essentially unchanged: 46% in the first test, and 47% in the second. The interaction (difference between the change in the challenged items and the change in the unchallenged items) was significant,  $t(28) = 2.80$ ,  $p = .0046$ .

The lack of change in the unchallenged items suggests that the challenge effect did not transfer to these items. It seems that subjects did not think of counterexamples unless they were asked to do so.

The effect of the challenge on PVs can also be found in between-subject comparisons. We formed two artificial groups of subjects. Subjects in the new challenged group were those in the original all-challenged group plus the scores for the challenged items in the half-challenged group. Subjects in the new unchallenged group were those in the original unchallenged group plus the scores for the unchallenged items in the half-challenged group. The proportion of PVs declined by 14% in the new challenged group and increased by 2% in the new unchallenged group,  $t(90) = 3.50$ ,  $p = .0004$ .

We also found a challenge effect on threshold ratios. Within the half-challenged group, the increase in thresholds was greater in the challenged items than in the unchallenged items,  $t(27) = 1.86$ ,  $p = .0367$ . In a between-group analysis, analogous to that for PVs just described, the new challenged group showed a mean threshold increase of .025 and the new unchallenged group showed a decrease of .022,  $t(88) = 1.98$ ,  $p = .0254$ . In summary, we conclude that thresholds increased as a result of challenge.

We classified answers to the counterexample questions (without knowledge of any other responses) into the following categories: no-trade—cannot think of any benefits great enough to justify the action (or provides a reason why such trade-offs are never required); in-kind trade—willing to give up  $X$  in order to save more  $X$  (e.g., species for species); out-of-kind trade—willing to give up one thing for something else that is mentioned (e.g., species for human lives); and unspecified trade—expresses general willingness to make a trade-off without specifying the benefit, denies that the harm is serious (e.g., in the discrimination item [5]), or provides some condition under which a trade can be made, without specifying the benefit (e.g., if other ways of avoiding the harm are impossible). Table 3 shows the numbers of these responses, classified according to initial and final PV responses, summed across all subjects.

When PVs were initially absent, it is not surprising that subjects were almost always willing to trade off the outcome in question. When PVs were present, reflection did lead to counterexamples in a total of 34 out of 51 cases (rows 3 and 4 of Table 3). Thinking of counterexamples led to change in the PV response in 17 cases (row 3), but, in 16 cases (row 4: 34–18), the PV response did not change despite the counterexample. Most of the counterexamples were in-kind: People would trade off lives for lives or species for species. Of course, such trade-offs were exactly what our omission items tested, therefore these would be sufficient for change in threshold. Although the prevalence of in-kind counterexamples would be predicted by other results (Beattie & Baron, 1995), the number of such counterexamples here may also have been affected by our giving the omission items first.

Table 3  
*Number of Responses to Challenge Manipulation, Classified According to Initial and Final Protected Value (PV) Responses, Summed Across all Subjects*

PV response		Response to challenge				Total
First	Second	No-trade	In-kind	Out-of-kind	Unspecified	
No	No	1	29	24	2	56
No	Yes	0	3	0	1	4
Yes	No	0	9	6	2	17
Yes	Yes	18	11	3	2	34
Total		19	52	33	7	111

Note. A response of "yes" means that the PV is present.

### Experiment 5

Experiment 5 examined explicit conflict between PVs. We compared resolution of conflict between two PVs with resolution of conflict between two NPVs. Would subjects be more or less likely to count both values when two PVs conflicted? The conflicts were between cutting government programs, with each program leading to sacrifice of some value. People may feel that the only way to resolve conflicts between PVs is to cut some from both or to cut only one. We also asked whether making a cut would be morally wrong. People faced with a conflict between two PVs might feel that whatever they do is morally wrong. Finally, we asked whether subjects believed that each conflict could occur in the real world. Subjects may engage in wishful thinking, feeling that conflict could be avoided when conflict was more extreme. (Jervis, 1976, calls this "belief overkill.") Of course, they could also be correct that these conflicts occur less often than other conflicts.

### Method

Fifty-two subjects completed a questionnaire on the World Wide Web for \$5.00. Their median age was 23 years (range = 14–50); 69% were female, and 62% were students. The questionnaire began as follows:

Imagine you are the President of the U.S. But do not imagine that you would have to become a different person, with different values, in order to get elected. Answer the questions according to your own values as they are now.

Politicians, however, are often faced with unpleasant choices when they negotiate with other politicians (such as those in the U.S. Congress). If they give in on one thing, they can get something else in return. Sometimes the thing they get can be much greater, in terms of money, than what they give up.

This questionnaire has two parts. In the first part, you evaluate ten cuts in various programs. These are parts of deals that would save the government money. Again, the cost savings could be much greater than the cost of the programs themselves. The savings could be used for anything you value, whether it is cutting taxes, reducing the deficit, increasing funds for other programs, or some combination of these.

In the second part, you will answer similar questions involving tradeoffs between one program and another.

The 10 programs were (with percent PVs after each) as follows:

1. Preventing loss of endangered plant and animal species. (61)
2. Emergency famine relief in Africa. (46)
3. Enforcement of civil rights laws against racial discrimination. (64)
4. Enforcement of laws against clear-cutting trees in old-growth rain forests. (55)
5. Medical aid to the poor for life-saving cancer surgery. (68)
6. Recording Native American languages and stories before they are lost forever. (54)
7. Preventing decay of paintings in the National Gallery of Art. (43)
8. Preventing decay of original national documents, such as the Federalist papers. (48)
9. Enforcement of laws against sweatshops. (64)
10. Enforcement of laws against fishing in a way that kills dolphins painfully. (71)

After each item, subjects wrote all the letters that applied from the following list (with percent endorsements in parentheses after each):

- A. This cut is acceptable if the cost savings are great enough. (26)
- U. I cannot imagine any case in which this cut is acceptable. (29)
- V. Cutting this is immoral, whatever the consequences. (25)
- W. I can imagine situations in which the savings are great enough to justify this, but these situations do not happen in the real world. (8)
- X. There may be situations in the real world in which the savings are great enough, but people cannot recognize these situations, so it is best never to allow cuts like this. (15)
- Y. Even if there are situations where cutting programs like this is acceptable, doing this can lead down a slippery slope, which will cause people to do this even when it is not acceptable. (23)
- Z. This kind of cut is unacceptable as a general rule, but we should make exceptions to it if we are sure enough. (30)

PVs were defined as endorsement of any answer from U to Y, and subjects were not permitted to endorse one of these and answer A for the same item. (This definition is more inclusive than in other experiments to maximize the chance of finding conflicting PVs. Although few items would change as a result of a less inclusive measure, the less inclusive one is probably better for most purposes.) The next part began as follows:

The situations in this part concern conflicts between cuts in two different programs. Again, these cuts arise from deals that you must make with other politicians, so the sizes of the cuts may vary. But all cuts will have significant effects on the programs in question.

The subject was then presented with 20 pairs, chosen so that both members of the pair were either protected (Answers U–Y above) or not (Answer A), with the pairs chosen so as to come as close as possible to half of each kind (and otherwise in order from the beginning of the list). After each pair, the subject answered the following:

- A. I would cut #1 if I had to choose one.
- B. I would cut #2 if I had to choose one.

- C. I would prefer to cut some from both, rather than cut one but not the other.
- D. It would be morally wrong to cut #1, even for the purpose of avoiding a cut in #2.
- E. It would be morally wrong to cut #2, even for the purpose of avoiding a cut in #1.
- F. Choices like this can always be avoided.

It is important to note that subjects were not required to choose from Items A–C.

## Results

Four subjects had no PVs and 6 had 10 (out of 10); these subjects, and the 10 others with only one response of each type, were not usable for within-subject comparison of PV and NPV responses, leaving 32. The mean number of PV responses for this group was 5.5. The number of PVs was unrelated to the tendency to resolve conflict in different ways.

In the questions about conflict resolution, PVs and NPVs were similar and did not differ significantly in the proportion of cases in which the subject thought that one value should be given priority (53% for PVs, 57% for NPVs) or that some cuts should be made from both (35% for PV, 45% for NPV).

The average number of cases in which the subject said it would be morally wrong to cut one value to avoid a cut in the other was greater for PVs (34%) than for NPVs (15%),  $t(31) = 3.51$ ,  $p = .0007$ . In an average of 7% of the PV cases, the subject said that cutting either of the two conflicting values would be wrong, versus 0% for NPVs,  $t(31) = 2.07$ ,  $p = .023$ . Of most importance, in an average of 66% of the cases, subjects thought that it was not immoral to sacrifice even one of the two PVs. Thus, for many PVs, problems of potential conflict with other PVs are not insurmountable, at least for conflicts within people who held the same PVs.

In an average of 23% of the cases for PVs and 12% for NPVs,  $t(31) = 2.22$ ,  $p = .0171$ , the subject said that "choices like this can always be avoided." This suggests that subjects tend to think that PVs can be honored without serious conflict.

## Experiment 6

When people say that they will not trade off life for money, a common rejoinder is of the form, "So you would not cross the street—risking being killed by a car—to collect \$1,000 that you had just won?" People seem not to think of the possibility of diluting the value in question (here, life) by making it probabilistic rather than certain. If, however, they take seriously their commitment to value  $X$  being infinitely more important than  $Y$ , then this commitment should not change when the loss of  $X$  is only a risk rather than a certainty. Conversely, if they are willing to trade off  $X$  and  $Y$  when the probability of loss on  $X$  is low, then they should not become unwilling when the probability of loss is high, unless the utility of  $Y$  has a maximum.

This is even easier to argue for when the trade-off is a matter of policy because a policy, applied over and over, turns a probability into a frequency. Thus, a 10% chance of losing a species as a result of making a certain kind of repeated decision will, as the number of decisions increase, result in roughly 10% of the decisions leading to species loss, and some species loss will be nearly certain with enough decisions. So, if some species loss is compensated by

no amount of benefit, then a policy of taking a 10% chance on species loss should also be impossible to compensate.

Experiment 6 examined PVs for decisions to allow genetically engineered wheat as a function of the harm, the benefit, and the probability of each. The question is whether PVs are sensitive to probability of harm and benefit. The present questions concern exactly the situation in which PVs are defined, specifically those in which some harm is accepted to create some benefit.

Baron and Spranca (1997, Experiment 5) examined willingness to pay to save endangered species as a function of the probability of success (and other variables). Subjects were less sensitive to probability when they had PVs for endangered species than when they did not, but they were still somewhat sensitive. DeKay and McClelland (1996) also found effects of probabilities on values for endangered species preservation, but they did not examine PVs.

## Method

Forty-nine subjects completed a questionnaire on the World Wide Web for \$3.00. Their median age was 24 years (range = 17–48); 57% were female, and 61% were students. The questionnaire, called "New Wheat," began as follows:

Imagine you are a regulatory official of the U.S. government. Your task is to make a recommendation about whether a new type of wheat should be approved for use. Your recommendation will almost certainly be followed.

The new wheat is created by genetic engineering. This uses biological methods to change the genetic makeup of the wheat, much in the same way that breeding does, but genetic engineering is faster and can insert genes that cannot normally be inserted by breeding.

In each of the following cases, you are given information about the possible harm, the possible benefit, and the probability of each. Suppose that the estimates of probability come from the most objective source you can imagine (e.g., studies reported in reputable scientific journals done by scientists with no financial interests in the issue, or better than that, if you can imagine it).

If the benefit or harm does not happen, you can assume that there will be no benefit, or no harm. Thus, if the probability of a harm is 1 in 100, this means there is a 1% chance of that happening and a 99% chance of no harm at all.

In each case you are asked to indicate how likely you are to approve the wheat under these conditions. The scale goes from 1 (definitely not approve) to 5 (definitely approve); a response of 3 means "not sure," but please use that as little as possible.

Each of the 12 cases that followed involved one of the following three benefits combined with one of the four harms. Each case required two screens. The order of the 12 (3 × 4) cases was randomized for each subject.

*Benefit 1:* The cost of growing wheat will decrease in the U.S., so that farmers will make more profit and still sell the wheat at a lower price.

*Benefit 2:* The price of wheat will decrease worldwide, and more wheat will be grown, leading to increased supplies for countries with grain shortages.

*Benefit 3:* The new wheat will be grown in the world's poorest countries and will cut the rate of child malnutrition in half.

*Harm 1:* One out of every million people who eat the new wheat will get a stomach ache from an allergic reaction to the new genes.

*Harm 2:* One out of every million people who eat the new wheat will die from an allergic reaction to the new genes.



*Harm 3:* The genes will escape into the environment and cause a related wild plant species to become extinct.

*Harm 4:* The genes will escape into the environment and cause several species of wild plants and the animals that eat them to become extinct.

Following presentation of the benefit and harm for the case, the subject responded to 16 combinations of probability of benefit and probability of harm. The first 4 had a "1 in 10,000,000" chance of harm. The benefit probabilities were, respectively, 100%, 90%, 50%, and 10%, and the same probabilities were used for each group of four, in this order. The remaining groups of four had harm probabilities of 1 in 100,000; 1 in 1,000; and 1 in 10. The subject responded to each case on a scale from 1 (*no for sure*) to 5 (*yes for sure*). All this information was on one screen, including the harm and benefit. A JavaScript program asked subjects to redo the answers on the screen (with their old responses in view) if they gave a higher rating to the worse of two options (in terms of probability of harm or benefit).

The second screen associated with each case repeated the description of the harm and benefit and then asked the subjects to "write all the letters that apply" of the following:

- A. This is acceptable if the probability of harm is low enough. (65%)
- U. This is acceptable only if there is no harm. (19%)
- V. This is not acceptable even if there is no harm. (4%)
- W. Risking such harm is always immoral. (17%)
- X. This is immoral even if there is no chance of harm. (2%)
- Y. This would be acceptable if we are certain of no harm, but we could never be certain, so it is best never to allow this. (16%)
- Z. Even if there are situations where this is acceptable, doing this can lead down a slippery slope, which will cause people to do this even when it is not acceptable. (25%)

Of these items, endorsement of U or V was taken to define PVs. We also used a strict definition of PVs that required endorsement of U or V and endorsement of W or X, thus combining acceptability and morality items. The numbers in parentheses indicate the overall percentage of endorsement of each item. A JavaScript routine prevented subjects from endorsing U, V, and Y when A was also endorsed.

**Results**

The frequency of endorsing PVs depended on both harm and benefit. In a repeated measures analysis of variance of frequency of PV endorsement across the 12 harm-benefit pairs, the effect of harm and benefit were both significant,  $F(3, 144) = 8.57, p < .0001$ , for harm;  $F(2, 96) = 8.30, p = .0005$ , for benefit; the interaction was not significant. The mean percentage of PV responses was 11%, 29%, 18%, and 33% for the four types of harm, respectively, and 30%, 19%, and 19% for the three types of benefit, respectively. Thirty-nine percent of the subjects endorsed no PVs at all, and the median subject endorsed 2 (out of a possible 12). For the strict definition of PVs, only the effect of harm was significant,  $F(3, 144) = 4.01, p = .0090$ . The mean percentage of strict PV responses for the four types of harm were, respectively, 5%, 12%, 9%, and 14%. (Interactions were not tested because they would not be meaningful; see Loftus, 1978.)

Subjects endorsed actions even when they expressed PVs for the harm-benefit pair if the probability of benefit was high enough and the probability of harm low enough. Table 4 shows the proportion of cases in which subjects endorsed the wheat, as a function of these probabilities, for cases in which they had PVs, strict PVs, or no PVs. Of interest is that these endorsements were sensitive to both probabilities of benefit and harm for the PV cases ( $p < .01$  for the linear trend for both benefit and harm, for PV;  $p <$

**Table 4**  
*Proportion of Cases in Which People Endorsed Actions, as a Function of Whether or not They Endorsed Protected Values (PVs) and of the Probability of Harm and Benefit*

Harm probability	Benefit probability			
	1.00	0.90	0.50	0.10
Cases with no PVs				
1/10,000,000	.855	.818	.593	.490
1/100,000	.724	.652	.458	.375
1/1,000	.506	.412	.272	.229
1/10	.337	.257	.174	.134
Cases with PVs				
1/10,000,000	.430	.342	.154	.134
1/100,000	.282	.222	.092	.075
1/1,000	.096	.046	.008	.008
1/10	.075	.042	.008	.008
Cases with strict PVs				
1/10,000,000	.382	.255	.095	.065
1/100,000	.245	.128	.000	.000
1/1,000	.167	.083	.000	.000
1/10	.167	.083	.000	.000

*Note.* The basic definition of PVs was agreement with U or V. The strict definition was the basic definition and agreement with W or X.

.05 for strict PVs, based on 12 subjects). Of course, the tendency to act was also sensitive to probabilities in the cases without PVs. In summary, subjects who endorsed PVs were willing to accept actions that violate those PVs if the probability of harm was sufficiently low, the probability of benefit was sufficiently high, or the type of benefit was considered important.

**Experiment 7**

The result of Experiment 6 suggests that people give up their PVs when the probability of harm is low. This concession appears to be inconsistent with the expression of PVs as absolute values. However, if we are to conclude that subjects are inconsistent, we must assume that they are willing to apply expected-utility theory to their PVs. That is, we must assume that a risk of some outcome with infinite disutility has infinite disutility itself and should therefore be protected. Perhaps subjects are unwilling to make this inference. In other words, the apparent inconsistency observed in Experiment 6 may lie in people's use of probability rather than in the way they think about values. They may, for example, consider very low probabilities as essentially zero.

Experiment 7 examined the effects of amount of the harm as well as probability. In particular, it asked whether those with PVs are less sensitive to changes in amount of harm than to changes in probability. PVs as expressed should not be sensitive to the amount of harm at all. Thus, we would expect that the amount of harm would have no effect on PVs, even if the risk of harm did have an effect. More generally, we would expect that amount would have more effect on actions that violate NPVs than on those that violate PVs, whereas probability would affect both kinds of action.

## Method

Fifty-two subjects completed a questionnaire on the World Wide Web for pay. Their median age was 25 years (range = 13–65); 63% were female, and 52% were students. The questionnaire, called "Policies," began as follows:

Sometimes government policies have bad side effects. This questionnaire is about how you evaluate the side effects. In particular, which side effects should be permitted if the program that leads to them is beneficial enough. Some items concern the risk of side effects.

Ten issues were used in all questions: extinction of an endangered animal species, extinction of an endangered plant species, an instance of illegal racial discrimination, clear-cutting of a square mile of trees in an old-growth rain forest, refusing life-saving cancer surgery to someone who cannot pay, loss of a Native American language before it is recorded for posterity, irreversible decay of a painting in an important art museum, irreversible decay of an original national document, employment of a 10-year-old in a garment factory for 60 hours/week, and painful death of a dolphin in a net to catch tuna. The questionnaire had two parts, each concerning all 10 issues. The order of the parts was balanced. One part asked about PVs. The subject indicated which of the following (presented simultaneously) applied to each issue:

- A. This should be allowed if the probability of harm is low enough.
- B. This should be allowed if the amount of harm is low enough.
- C. This should be allowed even if there is a great deal of harm.
- U. This should be allowed only if there is no chance of any harm.
- V. Risking such harm is always immoral.
- W. This should be allowed if we are certain of no harm, but we could never be certain, so we should not allow this.
- X. Even if there are situations where this is acceptable, allowing it can lead down a slippery slope, . . . so we should not allow it.

Endorsement of A, B, or C together with U, V, W, or X was not allowed. PVs were defined as endorsement of U or W, and NPVs were defined as endorsement of A, B, or C.

The other part consisted of a table showing the probability of harm, the amount of harm in number of occurrences, and a space for the subject's response to the question of whether the harm should be allowed, on a scale from 1 (*no for sure*) to 5 (*yes for sure*). There was one table for each of the 10 issues. The table had 16 rows, in a fixed order. Successive groups of four had risks of 1 in 10,000,000; 1 in 100,000; 1 in 1,000; and 1 in 10, respectively. Within each group, the number of occurrences was 1, 5, 25, and 125, respectively. Thus, the first row asked whether the subject would ever take a 1 in 10,000,000 chance of a single occurrence. The subject was not allowed to give higher ratings to worse cases (with higher probability or greater number). Ratings of 4 or 5 were counted as "yes" responses, and the analysis was conducted on the number of "yes" responses only. (Many subjects used only 1 and 5.)

## Results

Figure 1 shows the main results. The mean frequency of accepting action is plotted as a function of probability of harm and amount of harm, separately for PVs and NPVs, for the 29 participants, who had at least one value in each category. (Of the remaining 23 participants, 11 had no NPVs, 11 had no PVs, and 2 had neither PVs nor NPVs.)

Action ("yes" responses) was more likely for NPVs than for PVs,  $t(28) = 4.41, p < .0001$ , for subjects who had both;  $t(20) = 2.34, p = .0149$ , for comparison of the two groups with only PVs or only NPVs, with  $M_s = .12$  and  $.34$ , respectively. However, action was sometimes endorsed for PVs when the probability or amount of harm was low. Of importance, both PV and

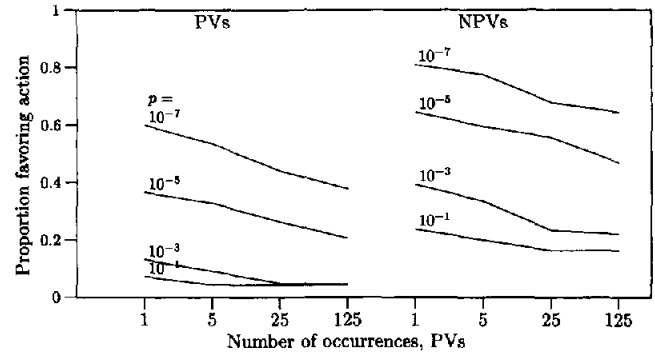


Figure 1. Proportion of "yes" responses (to accept action) as a function of amount of harm (1 to 125 occurrences) and probability of harm (separate lines), shown separately for protected values (PVs) and nonprotected values (NPVs), for those subjects who had some of each.

NPV responses were affected by amount as well as probability of harm. PV and NPV categories did not differ in the effect of probability (proportion of actions endorsed for lowest minus for highest probability, either within the group with both PV and NPV or between the two groups with only one type of response) nor in the effect of amount (both  $t$  tests nonsignificant). Subjects with PVs endorsed action more than half of the time when the probability was 1 in 10,000,000 of one instance of harm.

In a regression of the mean proportion of "yes" responses (concerning action) for PVs on age, sex, student status, and order, sex and order were significant (and the regression was significant overall). By using separate  $t$  tests, we found that males with PVs were more likely to endorse action than were females ( $M_s = .34$  vs.  $.13$ ),  $t(37) = 2.86, p = .0034$ . Action was more acceptable when the action ratings came first ( $M_s = .36$  vs.  $.18$ ),  $t(51) = 2.89, p = .0029$ . Neither result was significant for NPVs. It is difficult to interpret either of these effects. The order effect is of potential interest. Subjects may have tried to be consistent, but they did this mostly when they responded first to questions about values. Perhaps they had more trouble remembering the ratings they assigned.

## Discussion

PVs often yield to challenges. People who hold them can sometimes think of counterexamples. They will accept actions that violate PVs if the probability or amount of the harm is small relative to the probability and magnitude of benefit.

Some PVs may be simplified expressions of more complex beliefs, but relatively few subjects took the opportunity to endorse such statements as "I can imagine situations in which the benefits are great enough to justify this, but these situations do not happen in the real world" or "There are situations in the real world in which the benefits are great enough, but people can not recognize these situations, so it is best never to do this." Still, additional response options such as these may have affected the rate of endorsement of stronger options. In future research, the purpose of the research should dictate how many of these additional options are used. Too many of them may overwhelm the subjects.

People are reluctant to believe that PVs can conflict. When people are faced with such conflicts, however, most of the time they think that it is not immoral to simply resolve the conflicts in

some way, either by favoring one value or by making some compromise, in much the same way in which they make other trade-offs. Perhaps people understand this possibility even when they are faced with conflicts between their own PVs and those of others. This possibility requires further investigation.

Although people claim that amount of the violation of a PV does not matter (Baron & Spranca, 1997), PVs are less likely to be invoked if the amount of the violation is small or if it is improbable. This fact allows measurement of trade-offs. For the purpose of eliciting values, we can ask how people make trade-offs with small amounts, and we can extrapolate to larger amounts (as suggested by Schelling, 1968). A remaining question is whether people would find such extrapolation acceptable.

We did not attempt to combine all the ways of convincing people that their PVs are overgeneralizations. Our results suggest that this may be possible for the majority of PVs. The difficulty posed by PVs for social decision making may yield to a kind of cognitive therapy. The therapy would involve asking for counter-examples, asking people to imagine conflicts with other PVs, and asking people to imagine small amounts and low probabilities of harm. Such probing may be useful in negotiations that involve apparent PVs on both sides.

The case of small amounts and probabilities could also be used to elicit explicit trade-offs with NPVs. Once people see that such trade-offs can be made when probabilities are low, they may be more willing to engage themselves more seriously in the task of thinking about trade-offs, so that their values can be counted in social decision making along with the values of others who find trade-offs more natural.

Our results suggest that PVs are strong opinions, weakly held. They are strong in the sense that they express infinite trade-offs. Holders of these values assert that they are so important that they should not be traded off for anything. This assertion yields to a variety of challenges. After yielding, of course, the value may still be strong in the sense that a large amount of benefit is required to sacrifice the value.

The results are of greatest relevance to elicitation of values for public policy through the use of surveys. The results suggest that expressions of infinite trade-offs need not be accepted at face value and that respondents will change their expressions on probing. It remains to be determined whether probing will always suffice to elicit usable responses. Additional probes, other than those used here, may be needed. For example, it may be helpful to focus respondents on consequences, independently of the actions that produced them. (They could, for example, imagine that the consequences were not intended or caused by natural events.) The most general conclusion of the present studies is that researchers need to ask these questions before giving up on value elicitation for public policy.

The results may also be relevant to negotiation. Negotiation, however, differs from the present context in that exaggeration is almost expected at the outset, so that expressions of "nonnegotiable demands" may be even more subject to change than the PVs in our experiments. Still, some of these expressions may not be exaggeration. They may be initially honest expressions that may still yield to further probing.

Our results are not relevant to the prediction or understanding of actual behavior. People who hold PVs for endangered species, for example, may behave in ways that contribute to the destruction of

these species, such as buying products whose production pollutes the environment. People may do this even knowingly, and they may rationalize it in various ways when they are forced to become aware of it. The fact that people violate their own expressed values, however, need not imply that the expressed values themselves are insincere. When people truly confront the consistency of their values and their behavior, without the dulling effect of rationalization, they may feel guilty, and this feeling is a sign that the values are real even when they are not expressed in behavior. Our results suggest that the challenging of PVs may be a useful kind of cognitive therapy for such guilt feelings. The feelings may be unwarranted if they do not survive the challenge of such reflection.

## References

- Anderson, E. (1993). *Value in ethics and economics*. Cambridge, MA: Harvard University Press.
- Baron, J. (1973). Semantic components and conceptual development. *Cognition*, 2, 189–207.
- Baron, J. (1997). Biases in the quantitative measurement of values for public decisions. *Psychological Bulletin*, 122, 72–88.
- Baron, J. (1999). Utility maximization as a solution: Promise, difficulties, and impediments. *American Behavioral Scientist*, 42, 1301–1321.
- Baron, J. (2000). *Thinking and deciding* (3d ed.). New York: Cambridge University Press.
- Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior and Human Decision Processes*, 70, 1–16.
- Bazerman, M. H., Moore, D. A., & Gillespie, J. J. (1999). The human mind as a barrier to wiser environmental agreements. *American Behavioral Scientist*, 42, 1277–1300.
- Beattie, J., & Baron, J. (1995). In-kind and out-of-kind penalties: Preference and valuation. *Journal of Experimental Psychology: Applied*, 1, 136–151.
- Birnbaum, M. H. (Ed.). (2000). *Psychological experiments on the internet*. New York: Academic Press.
- Calfee, J., & Winston, C. (1998). The value of automobile travel time: Implications for congestion policy. *Journal of Public Economics*, 69, 83–102.
- DeKay, M. L., & McClelland, G. H. (1996). Probability and utility components of endangered species preservation programs. *Journal of Experimental Psychology: Applied*, 2, 60–83.
- Gregory, R., Lichtenstein, S., & Slovic, P. (1993). Valuing environmental resources: A constructive approach. *Journal of Risk and Uncertainty*, 7, 177–197.
- Hare, R. M. (1981). *Moral thinking: Its levels, method and point*. Oxford, England: Oxford University Press.
- Jervis, R. (1976). *Perception and misperception in international politics*. Princeton, NJ: Princeton University Press.
- Keeney, R. L., & Raiffa, H. (1993). *Decisions with multiple objectives: Preference and value tradeoffs*. New York: Cambridge University Press.
- Lim, C. S., & Baron, J. (1997). *Protected values in Malaysia, Singapore, and the United States*. Unpublished manuscript, University of Pennsylvania.
- Loftus, G. R. (1978). On interpretation of interactions. *Memory & Cognition*, 6, 312–319.
- Louviere, J. J. (1988). *Analyzing individual decision making: Metric conjoint analysis*. Newbury Park, CA: Sage.
- Maio, G. R., & Olson, J. M. (1998). Values as truisms: Evidence and implications. *Journal of Personality and Social Psychology*, 74, 294–311.
- McClosky, H., & Brill, A. (1983). *Dimensions of tolerance: What Americans believe about civil liberties*. New York: Russell Sage Foundation.

- Mitchell, R. C., & Carson, R. T. (1989). *Using surveys to value public goods: The contingent valuation method*. Washington, DC: Resources for the Future.
- Portney, P. R. (1994). The contingent valuation debate: Why economists should care. *Journal of Economic Perspectives*, 8, 3-17.
- Ritov, I., & Baron, J. (1990). Reluctance to vaccinate: Omission bias and ambiguity. *Journal of Behavioral Decision Making*, 3, 263-277.
- Ritov, I., & Baron, J. (1999). Protected values and omission bias. *Organizational Behavior and Human Decision Processes*, 79, 79-94.
- Rokeach, M. (1973). *The nature of human values*. New York: Free Press.
- Schelling, T. C. (1968). The life you save may be your own. In S. B. Chace, Jr. (Ed.), *Problems in public expenditure analysis* (pp. 127-175). Washington, DC: Brookings Institution.
- Seligman, C., & Katz, A. N. (1996). The dynamics of value systems. In C. Seligman, J. M. Olson, & M. P. Zanna (Eds.), *The psychology of values: The Ontario symposium* (Vol. 8, pp. 53-75). Mahwah, NJ: Erlbaum.
- Shafir, E., Diamond, P., & Tversky, A. (1997). Money illusion. *Quarterly Journal of Economics*, 112, 341-374.
- Tetlock, P. E., Lerner, J., & Peterson, R. (1996). Revising the value pluralism model: Incorporating social content and context postulates. In C. Seligman, J. Olson, & M. Zanna (Eds.), *The psychology of values: The Ontario symposium* (Vol. 8, pp. 25-51). Hillsdale, NJ: Erlbaum.
- Thompson, L. L., & Gonzalez, R. (1997). Environmental disputes: Competition for scarce resources and clashing of values. In M. H. Bazerman, D. M. Messick, A. E. Tenbrunsel, & K. A. Wade-Benzoni (Eds.), *Environment, ethics and behavior* (pp. 75-104). San Francisco: New Lexington Press.

Received June 7, 1999

Revision received November 19, 1999

Accepted December 17, 1999 ■

### Call for Nominations

The Publications and Communications Board has opened nominations for the editorships of *Journal of Applied Psychology*, *Journal of Consulting and Clinical Psychology*, *Journal of Educational Psychology*, *Psychological Bulletin*, and *Journal of Personality and Social Psychology: Interpersonal Relations and Group Processes* for the years 2003-2008. Kevin R. Murphy, PhD, Philip C. Kendall, PhD, Michael Pressley, PhD, Nancy Eisenberg, PhD, and Chester A. Insko, PhD, respectively, are the incumbent editors.

Candidates should be members of APA and should be available to start receiving manuscripts in early 2002 to prepare for issues published in 2003. Please note that the P&C Board encourages participation by members of underrepresented groups in the publication process and would particularly welcome such nominees. Self-nominations are also encouraged.

To nominate candidates, prepare a statement of one page or less in support of each candidate and send to

- Margaret B. Spencer, PhD, for the *Journal of Applied Psychology*
- Donna M. Gelfand, PhD, and Lucia Albino Gilbert, PhD, for the *Journal of Consulting and Clinical Psychology*
- Lauren B. Resnick, PhD, for the *Journal of Educational Psychology*
- Janet Shibley Hyde, PhD, and Randi C. Martin, PhD, for *Psychological Bulletin*
- Sara B. Kiesler, PhD, for *JPSP: Interpersonal Relations and Group Processes*

Address all nominations to the appropriate search committee at the following address:

[Name of journal] Search Committee  
 c/o Karen Sellman, P&C Board Search Liaison  
 Room 2004  
 American Psychological Association  
 750 First Street, NE  
 Washington, DC 20002-4242

The first review of nominations will begin December 11, 2000.