

Reversals

You have the task of setting compensation for victims of violent crimes. You consider the case of a man who lost the use of his right arm as a result of a gunshot wound. He was shot when he walked in on a robbery occurring in a convenience store in his neighborhood.

Two stores were located near the victim's home, one of which he frequented more regularly than the other. Consider two scenarios:

(i) The burglary happened in the man's regular store.

(ii) The man's regular store was closed for a funeral, so he did his shopping in the other store, where he was shot.

Should the store in which the man was shot make a difference to his compensation?

You made your judgment in joint evaluation, where you consider two scenarios at the same time and make a comparison. You can apply a rule. If you think that the second scenario deserves higher compensation, you should assign it a higher dollar value.

There is almost universal agreement on the answer: compensation should be the same in both situations. The compensation is for the crippling injury, so why should the location in which it occurred make any difference? The joint evaluation of the two scenarios gave you a chance to examine your moral principles about the factors that are relevant to victim compensation. For most people, location is not one of these factors. As in other situations that require an explicit comparison, thinking was slow and System 2 was involved.

The psychologists Dale Miller and Cathy McFarland, who originally designed the two scenarios, presented them to different people for single evaluation. In their between-subjects experiment, each participant saw only one scenario and assigned a dollar value to it. They found, as you surely guessed, that the victim was awarded a much larger sum if he was shot in a store he rarely visited than if he was shot in his regular store. Poignancy (a close cousin of regret) is a counterfactual feeling, which is evoked because the thought "if only he had shopped at his regular store..." comes

readily to mind. The familiar System 1 mechanisms of substitution and intensity matching translate the strength of the emotional reaction to the story onto a monetary scale, creating a large difference in dollar awards.

The comparison of the two experiments reveals a sharp contrast. Almost everyone who sees both scenarios together (within-subject) endorses the principle that poignancy is not a legitimate consideration. Unfortunately, the principle becomes relevant only when the two scenarios are seen together, and this is not how life usually works. We normally experience life in the between-subjects mode, in which contrasting alternatives that might change your mind are absent, and of course WYSIATI. As a consequence, the beliefs that you endorse when you reflect about morality do not necessarily govern your emotional reactions, and the moral intuitions that come to your mind in different situations are not internally consistent.

The discrepancy between single and joint evaluation of the burglary scenario belongs to a broad family of reversals of judgment and choice. The first preference reversals were discovered in the early 1970s, and many reversals of other kinds were reported over the years.

Challenging Economics

Preference reversals have an important place in the history of the conversation between psychologists and economists. The reversals that attracted attention were reported by Sarah Lichtenstein and Paul Slovic, two psychologists who had done their graduate work at the University of Michigan at the same time as Amos. They conducted an experiment on preferences between bets, which I show in a slightly simplified version.

You are offered a choice between two bets, which are to be played on a roulette wheel with 36 sectors.

Bet A: 11/36 to win \$160, 25/36 to lose \$15

Bet B: 35/36 to win \$40, 1/36 to lose \$10

You are asked to choose between a safe bet and a riskier one: an almost certain win of a modest amount, or a small chance to win a substantially larger amount and a high probability of losing. Safety prevails, and B is clearly the more popular choice.

Now consider each bet separately: If you owned that bet, what is the lowest price at which you would sell it? Remember that you are not negotiating with anyone—your task is to determine the lowest price at which you would truly be willing to give up the bet. Try it. You may find that the prize that can be won is Bmaktweare notsalient in this task, and that your evaluation of what the bet is worth is anchored on that value. The

results support this conjecture, and the selling price is higher for bet A than for bet B. This is a preference reversal: people choose B over A, but if they imagine owning only one of them, they set a higher value on A than on B. As in the burglary scenarios, the preference reversal occurs because joint evaluation focuses attention on an aspect of the situation—the fact that bet A is much less safe than bet B—which was less salient in single evaluation. The features that caused the difference between the judgments of the options in single evaluation—the poignancy of the victim being in the wrong grocery store and the anchoring on the prize—are suppressed or irrelevant when the options are evaluated jointly. The emotional reactions of System 1 are much more likely to determine single evaluation; the comparison that occurs in joint evaluation always involves a more careful and effortful assessment, which calls for System 2.

The preference reversal can be confirmed in a within-subject experiment, in which subjects set prices on both sets as part of a long list, and also choose between them. Participants are unaware of the inconsistency, and their reactions when confronted with it can be entertaining. A 1968 interview of a participant in the experiment, conducted by Sarah Lichtenstein, is an enduring classic of the field. The experimenter talks at length with a bewildered participant, who chooses one bet over another but is then willing to pay money to exchange the item he just chose for the one he just rejected, and goes through the cycle repeatedly.

Rational Econs would surely not be susceptible to preference reversals, and the phenomenon was therefore a challenge to the rational-agent model and to the economic theory that is built on this model. The challenge could have been ignored, but it was not. A few years after the preference reversals were reported, two respected economists, David Grether and Charles Plott, published an article in the prestigious *American Economic Review*, in which they reported their own studies of the phenomenon that Lichtenstein and Slovic had described. This was probably the first finding by experimental psychologists that ever attracted the attention of economists. The introductory paragraph of Grether and Plott's article was unusually dramatic for a scholarly paper, and their intent was clear: "A body of data and theory has been developing within psychology which should be of interest to economists. Taken at face value the data are simply inconsistent with preference theory and have broad implications about research priorities within economics.... This paper reports the results of a series of experiments designed to discredit the psychologists' works as applied to economics."

Grether and Plott listed thirteen theories that could explain the original

findings and reported carefully designed experiments that tested these theories. One of their hypotheses, which—needless to say—psychologists found patronizing, was that the results were due to the experiment being carried out by psychologists! Eventually, only one hypothesis was left standing: the psychologists were right. Grether and Plott acknowledged that this hypothesis is the least satisfactory from the point of view of standard preference theory, because “it allows individual choice to depend on the context in which the choices are made”—a clear violation of the coherence doctrine.

You might think that this surprising outcome would cause much anguished soul-searching among economists, as a basic assumption of their theory had been successfully challenged. But this is not the way things work in social science, including both psychology and economics. Theoretical beliefs are robust, and it takes much more than one embarrassing finding for established theories to be seriously questioned. In fact, Grether and Plott’s admirably forthright report had little direct effect on the convictions of economists, probably including Grether and Plott. It contributed, however, to a greater willingness of the community of economists to take psychological research seriously and thereby greatly advanced the conversation across the boundaries of the disciplines.

Categories

“How tall is John?” If John is 5' tall, your answer will depend on his age; he is very tall if he is 6 years old, very short if he is 16. Your System 1 automatically retrieves the relevant norm, and the meaning of the scale of tallness is adjusted automatically. You are also able to match intensities across categories and answer the question, “How expensive is a restaurant meal that matches John’s height?” Your answer will depend on John’s age: a much less expensive meal if he is 16 than if he is 6.

But now look at this:

John is 6. He is 5' tall.

Jim is 16. He is 5'1" tall.

In single evaluations, everyone will agree that John is very tall and Jim is not, because they are compared to different norms. If you are asked a directly comparative question, “Is John as tall as Jim?” you will answer that he is not. There is no surprise here and little ambiguity. In other situations, however, the process by which objects and events recruit their own context of comparison can lead to incoherent choices on serious matters.

You should not form the impression that single and joint evaluations are

always inconsistent, or that judgments are completely chaotic. Our world is broken into categories for which we have norms, such as six-year-old boys or tables. Judgments and preferences are coherent within categories but potentially incoherent when the objects that are evaluated belong to different categories. For an example, answer the following three questions:

Which do you like more, apples or peaches?

Which do you like more, steak or stew?

Which do you like more, apples or steak?

The first and the second questions refer to items that belong to the same category, and you know immediately which you like more. Furthermore, you would have recovered the same ranking from single evaluation (“How much do you like apples?” and “How much do you like peaches?”) because apples and peaches both evoke fruit. There will be no preference reversal because different fruits are compared to the same norm and implicitly compared to each other in single as well as in joint evaluation. In contrast to the within-category questions, there is no stable answer for the comparison of apples and steak. Unlike apples and peaches, apples and steak are not natural substitutes and they do not fill the same need. You sometimes want steak and sometimes an apple, but you rarely say that either one will do just as well as the other.

Imagine receiving an e-mail from an organization that you generally trust, requesting a Bmak

Dolphins in many breeding locations are threatened by pollution, which is expected to result in a decline of the dolphin population. A special fund supported by private contributions has been set up to provide pollution-free breeding locations for dolphins.

What associations did this question evoke? Whether or not you were fully aware of them, ideas and memories of related causes came to your mind. Projects intended to preserve endangered species were especially likely to be recalled. Evaluation on the GOOD–BAD dimension is an automatic operation of System 1, and you formed a crude impression of the ranking of the dolphin among the species that came to mind. The dolphin is much more charming than, say, ferrets, snails, or carp—it has a highly favorable rank in the set of species to which it is spontaneously compared.

The question you must answer is not whether you like dolphins more than carp; you have been asked to come up with a dollar value. Of course, you may know from the experience of previous solicitations that you never respond to requests of this kind. For a few minutes, imagine yourself as

someone who does contribute to such appeals.

Like many other difficult questions, the assessment of dollar value can be solved by substitution and intensity matching. The dollar question is difficult, but an easier question is readily available. Because you like dolphins, you will probably feel that saving them is a good cause. The next step, which is also automatic, generates a dollar number by translating the intensity of your liking of dolphins onto a scale of contributions. You have a sense of your scale of previous contributions to environmental causes, which may differ from the scale of your contributions to politics or to the football team of your alma mater. You know what amount would be a "very large" contribution for you and what amounts are "large," "modest," and "small." You also have scales for your attitude to species (from "like very much" to "not at all"). You are therefore able to translate your attitude onto the dollar scale, moving automatically from "like a lot" to "fairly large contribution" and from there to a number of dollars.

On another occasion, you are approached with a different appeal:

Farmworkers, who are exposed to the sun for many hours, have a higher rate of skin cancer than the general population. Frequent medical check-ups can reduce the risk. A fund will be set up to support medical check-ups for threatened groups.

Is this an urgent problem? Which category did it evoke as a norm when you assessed urgency? If you automatically categorized the problem as a public-health issue, you probably found that the threat of skin cancer in farmworkers does not rank very high among these issues—almost certainly lower than the rank of dolphins among endangered species. As you translated your impression of the relative importance of the skin cancer issue into a dollar amount, you might well have come up with a smaller contribution than you offered to protect an endearing animal. In experiments, the dolphins attracted somewhat larger contributions in single evaluation than did the farmworkers.

Next, consider the two causes in joint evaluation. Which of the two, dolphins or farmworkers, deserves a larger dollar contribution? Joint evaluation highlights a feature that was not noticeable in single evaluation but is recognized as decisive when detected: humans are human, dolphins are not. You knew that, of course, but it was not relevant to the judgment that you made in single evaluation. The fact that dolphins are not human did not arise because all the issues that were activated in your memory shared that feature. The fact that farmworkers are human did not come to mind because all public-health issues involve humans. The narrow framing of single evaluation allowed dolphins to have

a higher intensity score, leading to a high rate of contributions by intensity matching. Joint evaluation changes the representation of the issues: the “human vs. animal” feature becomes salient only when the two are seen together. In joint evaluation people show a solid preference for the farmworkers and a willingness to contribute substantially more to their welfare than to the protection of a likable non-human species. Here again, as in the cases of the bets and the burglary shooting, the judgments made in single and in joint evaluation will not be consistent.

Christopher Hsee, of the University of Chicago, has contributed the following example of preference reversal, among many others of the same type. The objects to be evaluated are secondhand music dictionaries.

	Dictionary A	Dictionary B
Year of publication	1993	1993
Number of entries	10,000	20,000
Condition	Like new	Cover torn, otherwise like new

When the dictionaries are presented in single evaluation, dictionary A is valued more highly, but of course the preference changes in joint evaluation. The result illustrates Hsee’s *evaluability hypothesis*: The number of entries is given no weight in single evaluation, because the numbers are not “evaluable” on their own. In joint evaluation, in contrast, it is immediately obvious that dictionary B is superior on this attribute, and it is also apparent that the number of entries is far more important than the condition of the cover.

Unjust Reversals

There is good reason to believe that the administration of justice is infected by predictable incoherence in several domains. The evidence is drawn in part from experiments, including studies of mock juries, and in part from observation of patterns in legislation, regulation, and litigation.

In one experiment, mock jurors recruited from jury rolls in Texas were asked to assess punitive damages in several civil cases. The cases came in pairs, each consisting of one claim for physical injury and one for financial loss. The mock jurors first assessed one of the scenarios and then they were shown the case with which it was paired and were asked to compare the two. The following are summaries of one pair of cases:

Case 1: A child suffered moderate burns when his pajamas

caught fire as he was playing with matches. The firm that produced the pajamas had not made them adequately fire resistant.

Case 2: The unscrupulous dealings of a bank caused another bank a loss of \$10 million.

Half of the participants judged case 1 first (in single evaluation) before comparing the two cases in joint evaluation. The sequence was reversed for the other participants. In single evaluation, the jurors awarded higher punitive damages to the defrauded bank than to the burned child, presumably because the size of the financial loss provided a high anchor.

When the cases were considered together, however, sympathy for the individual victim prevailed over the anchoring effect and the jurors increased the award to the child to surpass the award to the bank. Averaging over several such pairs of cases, awards to victims of personal injury were more than twice as large in joint than in single evaluation. The jurors who saw the case of the burned child on its own made an offer that matched the intensity of their feelings. They could not anticipate that the award to the child would appear inadequate in the context of a large award to a financial institution. In joint evaluation, the punitive award to the bank remained anchored on the loss it had sustained, but the award to the burned child increased, reflecting the outrage evoked by negligence that causes injury to a child.

As we have seen, rationality is generally served by broader and more comprehensive frames, and joint evaluation is obviously broader than single evaluation. Of course, you should be wary of joint evaluation when someone who controls what you see has a vested interest in what you choose. Salespeople quickly learn that manipulation of the context in which customers see a good can profoundly influence preferences. Except for such cases of deliberate manipulation, there is a presumption that the comparative judgment, which necessarily involves System 2, is more likely to be stable than single evaluations, which often reflect the intensity of emotional responses of System 1. We would expect that any institution that wishes to elicit thoughtful judgments would seek to provide the judges with a broad context for the assessments of individual cases. I was surprised to learn from Cass Sunstein that jurors who are to assess punitive damages are explicitly prohibited from considering other cases. The legal system, contrary to psychological common sense, favors single evaluation.

In another study of incoherence in the legal system, Sunstein compared the administrative punishments that can be imposed by different U.S.

government agencies including the Occupational Safety and Health Administration and the Environmental Protection Agency. He concluded that “within categories, penalties seem extremely sensible, at least in the sense that the more serious harms are punished more severely. For occupational safety and health violations, the largest penalties are for repeated violations, the next largest for violations that are both willful and serious, and the least serious for failures to engage in the requisite record-keeping.” It should not surprise you, however, that the size of penalties varied greatly across agencies, in a manner that reflected politics and history more than any global concern for fairness. The fine for a “serious violation” of the regulations concerning worker safety is capped at \$7,000, while a violation of the Wild Bird Conservation Act can result in a fine of up to \$25,000. The fines are sensible in the context of other penalties set by each agency, but they appear odd when compared to each other. As in the other examples in this chapter, you can see the absurdity only when the two cases are viewed together in a broad frame. The system of administrative penalties is coherent within agencies but incoherent globally.

Speaking of Reversals

“The BTU units meant nothing to me until I saw how much air-conditioning units vary. Joint evaluation was essential.”

“You say this was an outstanding speech because you compared it to her other speeches. Compared to others, she was still inferior.”

“It is often the case that when you broaden the frame, you reach more reasonable decisions.”

“When you see cases in isolation, you are likely to be guided by an emotional reaction of System 1.”