

**Section 1**

(For BSF use Only)

Application no:

**Full title:**

Intuitions about punishment and deterrence in law: heuristics, biases, and the role of emotions

---

---

**Principal Investigators:**

Ilana Ritov, Hebrew University

Jonathan Baron, University of Pennsylvania

---

---

**Abstract:**

An abstract of about two hundred words should describe the (1) objectives, (2) methodology, and (3) significance of the proposed research.

***Abstract***

In our 1993 paper on Intuitions about penalties and compensation," we took the law-and-economics approach as roughly normative, and we pitted this against people's intuitions. Through a combination of heuristics (e.g., preference for nature) and mis-education (e.g., failure to understand incentive and deterrent effects of tort rules), people's intuitions did not match the system very well. This raised questions about how the system is maintained if so few people understand it, about whether the law in fact conforms to economic rationality, and, if not, about whether it can.

We propose here to extend our approach to other aspects of the law, in addition to tort law. Questions about deterrence and incentive run through almost all of the law, including criminal, tort, contract, and property. A particular emphasis will be given to the role of probability throughout several domains.

The general idea also fits with our previous work on emotion. Punitiveness is surely affected by emotion. But emotion might also interact with other factors, reducing sensitivity to relevant factors like probability or deterrability, and possibly increasing sensitivity to irrelevant factors (like, arguably, the extent to which the injurer is subject to envy). In as many of the studies as possible, we shall manipulate the emotional significance of cases, looking for interactions between that significance and other manipulations. We shall also include manipulation checks, of the sort we have been using in our current research.

Because U.S. and Israeli legal systems differ someone, we can also address, in a general way, questions about the effect of the legal environment on public beliefs about the law.

## Detailed Description of the Research Plan

Below is a numbered list of required information for this section. Please mark the headings accordingly.

1. Brief description of the subject and of the scientific and technological background
2. Objectives and expected significance of the research
3. Comprehensive description of the methodology and plan of operation
4. Detailed account of available U.S. and Israeli resources, including all personnel and equipment
5. One-page progress report on any recent BSF-supported project
6. Relevant bibliography on research area
7. Time schedule and work plan

### Note:

Items 1 through 4 are limited to a total of 15 pages, including figures.

### 1. **Brief description of the subject and of the scientific and technological background:**

In our 1993 paper on "Intuitions about penalties and compensation," we took the law-and-economics approach as roughly normative, and we pitted this against people's intuitions. Through a combination of heuristics (e.g., preference for nature) and mis-education (e.g., failure to understand incentive and deterrent effects of tort rules), people's intuitions did not match the system very well. This raised questions about how the system is maintained if so few people understand it, about whether the law in fact conforms to economic rationality, and, if not, about whether it can. These questions have now been picked up by others (often with due credit to our various contributions).

We propose here to extend our approach to other aspects of the law, in addition to tort law. Questions about deterrence and incentive run through almost all of the law, including criminal, tort, contract, and property.

We shall also examine the role of probability throughout several domains. For example, even in property law, private property provides an incentive to develop property, but that requires risk taking. (Various proposed patent-law reforms have actually tried to reduce the risk, e.g., by giving inventors prizes roughly equal to the expected value of their inventions, as discussed by Shavell, 2003.)

The general idea also fits with our previous work on emotion. Punitiveness is surely affected by emotion. But emotion might also interact with other factors, reducing sensitivity to relevant factors like probability or deterrability, and possibly increasing sensitivity to irrelevant factors (like, arguably, the extent to which the injurer is subject to envy). In as many of the studies as possible, we shall manipulate the emotional significance of cases, looking for interactions between that significance and other manipulations. We shall also include manipulation checks, of the sort we have been using in our current research.

Because U.S. and Israeli legal systems differ someone, we can also address, in a general way, questions about the effect of the legal environment on public beliefs about the law.

### 2. **Significance**

The studies we propose should add to our understanding of the nature of law and legal systems and their basis in human judgment. They may point out areas where public education in the function of law is needed. Baron and Ritov (1993), for example, found that very few subjects (including judges) were familiar with the idea that tort penalties were justified (at least in part) by deterrence. This is such a fundamental and common idea that we found this result quite surprising. Most of the experiments we have proposed could lead to similar results, or not. It will be of interest which features of the economic approach to law are understood and which are not.

The work will also contribute to the understanding of the role of emotion in people's thinking about legal and economic issues. Although this understanding need not have such direct implications as that concerning educational gaps, it is still helpful to those who want to understand how laws come to exist and where they get

their support.

A common assumption of "behavioral law and economics," which we share, is that much can be understood by looking at basic principles of human judgment. These principles govern the judgments of citizens, legislators, and those involved in making laws work (judges, police, lawyers). This is a different approach than (for example) the "public choice" perspective, which tends to think of people as rational actors and tries to understand anomalous laws through principles of collective action, pressure groups, and so on. That approach is valuable too, but it is probably not sufficient by itself (as argued by Brennan & Lomasky, 1993, and Green & Shapiro, 1994, among others).

### **3. Comprehensive description of the methodology and plan of operation**

#### **3.1 General method**

We shall carry out our research both with paper questionnaires (in Israel) and on the World Wide Web. Baron has been doing research on the web since 1996. For the last few years, he has been programming questionnaires with JavaScript, which presents questions on successive screens. Usually the screens represent conditions in a factorial design presented in a random order chosen for each subject. These studies can check answers and make later questions contingent on earlier answers. Current experiments are described at <http://www.psych.upenn.edu/~baron/qs.html>. We send email to about 500 subjects whenever a new study is ready, and we usually get 100 responses in 4 hours, at which point the study is removed. Sometimes we stop at 50. The results are monitored as they come in, and studies that seem to be misunderstood can be redesigned.

Use of the Web for research has several advantages over the alternatives for this kind of research (usually conducted with undergraduates): the subjects are much more varied than those from other convenience samples; expenses connected with data entry and checking are reduced; and, because it is easy to check answers as the subject enters them, fewer responses need to be discarded because they are nonsensical (Birnbaum, 2000). Moreover, the general quality of the data is at least as high as that of data from paper questionnaires, and, in general, substantive results do not differ from those of comparable methods (Birnbaum, 1999; McGraw et al., 2000). Because subjects are paid, it is possible to track individual identities to ensure that nobody completes the same study twice. Subjects are also timed without their knowledge so that data can be eliminated from subjects who complete the questionnaire too quickly to have paid attention to the questions and their responses. Finally, demographics of this population closely match the 2000 U.S. census on all demographic characteristics (income, education, race, and marital status) except gender (more females in the sample population than in the U.S. census; Babcock et al., 2003).

The general experimental procedure (which is continuing to evolve) is to use JavaScript to randomize the stimuli separately for each subject. Each session contains 20-100 judgments, depending on their difficulty, and pay ranges from \$3 to \$6, again, depending on the difficulty and time required. In some cases, stimuli are presented in blocks, and the blocks are randomized too. Typically, each item is presented on a screen, so that the subject cannot back up.

Most of the web experiments involve within-subject factorial designs, with the conditions presented in a different random order to each subject. We often manipulate variables of little interest, such as the number of dollars involved, simply to get more data from each subject without repeating cases identically. This procedure permits a within-subject analysis of effects. It is possible to find a significant effect in one direction while simultaneously finding some subjects with results that are significant in the opposite direction, even after correction for multiple tests. (For this purpose, we use the step-down re-sampling procedure of Westfall and Young, 1993, as implemented by Dudoit and Ge, 2003).

It is possible to do between-subject designs on the web, but most subjects expect to get paid a few dollars at a time. Thus, it is necessary to attach a smaller questionnaire as a "bonus" to other studies. For between-subject designs, we will rely largely on paper questionnaires done in Israel, where it is possible to find large groups of students who will complete short questionnaires.

#### **3.2 Punitive damages and criminal penalties**

The economic theory of punitive damages, which applies also to many criminal penalties, is quite simple: the purpose of punishment is to achieve socially optimal level of deterrence (Polinsky and Shavell, 1998). Compensation awards are sufficient for that purpose if all harms are detected and fully compensated. (But full compensation is often impossible, e.g., in cases of wrongful death, in which the dead person cannot be

compensated.) Punitive damages in tort cases should be awarded "if and only if an injurer has a chance of escaping liability for the harm he causes" (p. 874). In that case the punitive damages should be determined, based on the magnitude of the harm and the probability of detection. The same principle of probability is used in the economic analysis of criminal law.

Sunstein et al. (2000, henceforth SSK) argue that people do not attempt to achieve optimal deterrence, not only because they don't understand or appreciate the role of punitive damages in achieving optimal deterrence, but also because they simply don't agree with the principle that punishment should primarily be determined by the goal of achieving deterrence. These are two different questions. We refer first to the question of the deliberate rejection of the principle.

There are two aspects of to punitive damages. One is to penalize not only for bad outcomes, but also for bad behavior. This aspect is intuitively appealing, and will certainly be endorsed by most people. This purpose is made quite explicit in the jury instructions. Although the instructions to the jury state that punitive damages have a double role: to punish the defendant and to deter the defendant and others from committing similar actions in the future, with regard to punishment that instructions go on: "Punitive damages should be awarded if a preponderance of the evidence shows that the defendants acted either maliciously or with reckless disregard for the welfare of others" (p. 19, appendix of the SSK paper). So, in fact, subjects (or jurors) are explicitly asked to take into account the reprehensibility of the injurer. This is in contrast to the theory proposed by Polinsky and Shavell, who explain why reprehensibility should not affect damages. In any case, it is hardly surprising that given the explicit punitive role of punitive damage awards people, even Chicago law students (participants in the second experiment) reject the option not to set any punitive damages, on account of perfect detectability, when the injurer's behavior is described as particularly reckless.

### *3.2.1 Probability of detection*

The other aspect of punitive damages involves deterrence. Here, the role of detection probability should come into play. However, one could object to the role of this factor, not just because of a strong emotional response to the outcome or the behavior of the injurer, but also on normative grounds. In order to achieve optimal deterrence one needs to deviate from the principle that like cases should be sentenced alike.

We have already shown (Baron & Ritov, 1993) that many people do not understand the theory of deterrence and its implications for setting punitive damages. We propose to pursue the specific role of the judged probability of detection. This is not just a psychological question. The theory is not absolutely clear either. In theory the court should base the award on the post-hoc assessment of the probability that the harm in question would have been detected (assuming that the purpose of the awards is to deter others from causing the same harm). How should we define this probability. Obviously, as Polinsky and Shavell acknowledge, this probability must be related to the categorization of the harm. They recommend that, other things being equal, narrow definitions of the harm should be used.

There are several questions that can be asked. First, do people understand that probability of detection should matter. Proponents of punitive damages (for example Eisenberg et al., 1997) argue that punitive damages are "well behaved." They cite the positive correlation of punitive damages and compensation awards as important evidence supporting their claim. However, as Polinsky notes in response (1997) it is more likely that the theory would predict negative, rather than positive, correlation, because, as the level of compensatory damages rises, so would the probability of detection and liability.

SSK found that probability of detection had no effect on judgments of penalties. They suggest themselves that the results might be different if subjects thought that differences in probability of detection were the result of differences in the furtiveness of the crime. We shall explore this question by comparing the effects of various manipulations of probability. In the Sunstein et al. paper, probability changed apparently because of the function of the law itself, not the infraction. The theory of optimal deterrence does not change, but people might think of these effects differently.

SSK's major experiment showing that judgments are unaffected by probability of detection compared different cases. Subjects were presented with three cases differing in many details, including the probability of detection. The probability of detection did not have a significant effect on the damage awards. Although probability of detection was varied within subject, the subjects did not judge the same case with different probabilities. So the conclusion from this experiment should be that in separate judgment probability of detection is not a significant factor. In order to go a step further and claim that subjects deliberately reject its role, one needs to do a real within-subject experiment (as the authors acknowledge).

Another factor, not considered by Sunstein et al., is fairness. When probability of detection is not under the violator's control, it seems unfair to punish people differently because of a factor they cannot control, such as the number of law enforcement agents assigned to the type of infraction they committed. We shall test this by attempting to "debias" this intuition, in an explicit within-subject design where it is clearly apparent that like cases (from the violator's perspective) are not being treated alike. In particular, we shall point out that unfairness is already present to a large degree, because getting caught is a matter of luck. This manipulation is similar to one found to be effective by Ubel et al. (2001), which reduced the concern for between-group equality by pointing to other groups that were necessarily treated unequally.

We shall also test the hypothesis that the reluctance to break the "equal crime equal punishment" rule may be less intense if the deterrence applied mostly (or solely) to the offender himself. This situation would focus the subject on the offender and be less likely to evoke fairness considerations. We would have to make sure though that the punitive damages award is not interpreted as punishment for already committed (but undetected) harms. We shall also test this hypothesis by asking for separate judgments of "punitive penalties" and "deterrence penalties" (just as juries are now asked to separate compensation and punishment).

In our experiments, we shall use a number of examples as cases, which differ in their emotion-arousing properties. For example, cases that involve psychological injury to children will probably be more emotional than cases that involve damage to replaceable property. We shall check our manipulation by asking the subjects to rate emotionality, as we have done in our recent studies under our current grant. We shall test the hypothesis that emotionality reduces sensitivity to probability (if any exists at all). When the situation is emotionally arousing, subjects will more likely base their judgments on the consequences that arouse the emotion.

### *3.2.2 Cost of enforcement*

In the case of criminal penalties, the law enforcement system must make a trade-off between investing money in detection, to increase the probability, or using more severe penalties. Some times the penalties are costly, like prison. When the offenders can pay fines, the level of the fines may somewhat cover the costs of policing; but, the higher the fines, the fewer offenders capable of paying them, and the greater the requirement for penalties that are costly to the state. Thus, the money spent on detection should depend on the cost of penalties to the state. Crimes that could be punished either with prison or fines should receive enough funding on detection so as to minimize the use of prison.

We hypothesize that most subjects will favor spending more money on detection even when it is inefficient to do so. More generally, they will not consider the cost of enforcement or the cost of penalties (prison vs. fines). To test these hypotheses, we will manipulate costs of detection and costs of punishment and ask subjects to choose among approaches that differ in the trade-off between resources devoted to detection and severity (hence also cost) of the penalty. The trade-off will differ across cases. In some cases, the cost of punishment will rise steeply as the cost of detection decreases, and in other cases the rise will be much slower. We expect subjects to fail to choose optimally because they will be insufficiently sensitive to this trade-off.

Again, we shall manipulate the emotionality of cases. We expect emotionality should increase the choice of more severe punishment (rather than increase the effort at detection), in the trade-off. This is because the severity of the punishment is a dimension more compatible with the emotion in question, which we expect to be anger. Anger is associated with retaliation, once the perpetrator is in hand. It is not so clearly associated with the work of the detective.

### *3.2.3 Categorization*

As the probability that the harm in question would have been detected is dependent upon the categorization of that harm, we shall examine the categorization question more extensively. Categorization of an object is closely linked with the perceived possible variations of the object's features. Those variations are certainly affected by prior experience, but they are also constructed during the judgment process itself. This idea goes back to the Gestalt concept of pattern goodness, and is further developed in Garner's research on the processing of information and structure (1974). According to this theory, as a target is judged in the absence of an explicitly defined context set, an inferred set of stimuli, including the target stimulus, is constructed. The inferred set reflects one's encoding of the stimulus properties, based on the expectations regarding the possible variations in these properties. Garner's research also demonstrated the link between the goodness of a pattern and the size of the inferred subset of related patterns. Due to redundancy, better patterns evoke smaller sets. When subjects were presented, for example, with a set of 90 dot patterns, subjective "goodness" rating of the

patterns was negatively correlated with the size of the group in which each pattern was placed, in a free classification task.

Kahneman and Miller (1986) extended the notion of inferred set beyond the domain of perception, and argued that most judgments involve spontaneously evoked set of alternatives. Although in many cases the generation and use of the invoked alternatives involve conscious thought processes, often motivated by affective goals, in many other cases the alternatives are generated automatically and unconsciously (Kahneman, 1995).

Returning to our question concerning the impact of spontaneous categorization on the judged probability of detection, we suggest that this judgment must be affected by the evoked set of similar cases. The size of this set is likely to affect the judged probability. In particular, the psychological theory scantily described above would suggest that the more distinct the target case, the larger the evoked set, and the lower the perceived probability of detecting similar cases. We propose to test this prediction by providing additional unrelated details regarding the target case, and testing the effect of those details on the judged probability of detection.

On the other hand, Fox and Birke (2002) propose that a richer description yields a higher probability judgment. This mechanism would lead to the opposite prediction. We might expect both sorts of effects under different circumstances.

We might also expect emotionality to lead to a higher probability judgment. We shall manipulate emotionality in order to test this hypothesis.

In another study we shall address the same question by asking about the effect of a previous "context case", varying the degree to which the two cases are likely to be categorized together.

#### *3.2.4 Probability in hindsight*

A number of other questions involve the estimated probability of detection. Polinsky and Shavell state that if the injurer's subjective belief about the probability of being found liable is different from the "objective" probability, then, in principle, the subjective probability should be used in the calculation. Recent research by Tykocinski (2001) suggests that people adjust probabilities in hindsight in order to avoid disappointment, raising the probability of bad events after they happen so as to make them seem more expected. Thus, as the injurer has been found liable, his ex-post estimate of the probability is predicted to be too high. This estimate may not be applicable when the injurer, or other similar potential injurers must decide, ex-ante, what level of precautions they want to take.

Rachlinsky (1998) discusses the role of hindsight bias in assessing probability in the punitive damages case, although he does not target punitive damages in particular. The importance of determining, ex-post, the probability that a harm would have happened, when it is known that it has, is a difficult one. The hindsight bias effectively turns a negligence rule to a quasi-strict liability one. He argues that if the bias is severe enough the result is optimal, as in the case of strict liability. In other cases it results in incentive to take excess care, or, more likely, in an excessive reduction of activity level. (Obstetricians, for example, may revert to gynecology if they frequently lose lawsuits because of hindsight bias. The theoretical equivalence of strict liability and negligence assumes that activity level stays constant.) Rachlinsky acknowledges that there is no general remedy for the hindsight bias, but suggests that various practices provide at least partial solutions in specific cases.

Some of the measures he discusses, like suppression of evidence concerning subsequent remedial measures, and compliance with ex ante norms or standards, suggest that accidents resulting from acts and accidents resulting from inaction are treated differently. This is of course consistent with our work on omission bias. However in these experiments probabilities were always clearly specified (for a good reason). It might also be true that people would judge the probability of a harm as a result of omission to be lower. We shall thus compare the ex-ante and ex-post probabilities for omission and for action.

Hindsight bias may be greater for acts than for negligent omissions. In particular, subjects or jurors may put themselves in the position of the injurer. When the injurer causes harm through an action, it will be easy to suppose that the injurer thought about what would happen. When an omission is involved, however, the injurer just may not have thought about it at all. This would lead subjects to suppose that the injury was in fact not all that foreseeable.

These effects may be mediated by emotional responses. Omissions may be less emotion-arousing than acts.

### 3.2.5 *Potential harm*

Another question concerns the role of potential harm. Polinsky and Shavell argue that only the actual harm should enter the computation of punitive damages, because if potential harm is taken into account that will lead to excessive precautions. This requirement calls for judgment based only on outcome.

Earlier research (Baron and Hershey, 1988) showed that people are strongly affected by the outcome when evaluating the quality of the decision that led to the outcome, perhaps due to an over-generalization of the rule that bad decisions lead to bad outcomes. Still, this outcome bias does not mean that information concerning the behavior leading to the outcome will be completely ignored.

In fact, we might expect the opposite bias. When people are asked to judge on the basis of outcome only, they will still be affected by the riskiness of the behavior. This bias is not really opposite, however, since both biases are the result of attending to information that is usually relevant but is not supposed to be relevant to the case at hand.

The respective role of outcome and behavior will be examined by varying orthogonally the severity of the outcome and the reprehensibility of the injurer's behavior. We shall be careful, of course, to present cases in which the two factors are not directly linked. For example, the outcome of construction defects may be that a young person is injured and endures a broken leg, or that an old lady is injured and suffers from a heart attack as a result. We shall also ask about assessed probability of detection, to examine the possible mediation effect of this factor. And we shall ask about the probability of harm, for the same reason. We shall instruct subjects either to attend to outcome and ignore the behavior, attend to the behavior and ignore the outcome, or neither.

Although we shall vary emotionality, we have no hypothesis about whether it should make people attend more to the behavior or the outcome. Both are stimuli for anger.

### 3.2.6 *Risk aversion*

Polinsky and Shavell also suggest that, when juries are asked to come up with an estimate, they may err, but it doesn't matter, as long as they are not generally biased. However, risk aversion in either the defendant or the jury may have systematic effect on the probability estimates. For example, Polinsky and Shavell argue that punitive damages for risk averse injurer should be lower. This makes sense only if the potential injurers that are supposed to be deterred by the award are also risk averse.

We shall examine risk aversion using a "point of view" manipulation, of the sort used successfully by Birnbaum and Stegner (1979), in which subjects take the point of view of a potential injurer (e.g., a doctor faced with taking the time for additional tests that might reduce risk) and a potential victim (e.g., a patient). We shall also manipulate risk by varying the extremeness of the possible outcomes, their distance from the expected outcome. We shall also manipulate independently the risk (dispersion) of the magnitude of the injury and the risk of the monetary award for causing it. Prospect theory predicts an interaction between point of view and risk, but only for the monetary award. This is because the award is a loss for the injurer but (we assume) a gain for the victim.

We shall also ask whether subjects in the role of jurors are sensitive to the risk attitudes of the defendant, or the general norms concerning risk taking in the relevant industry or activity. Penalties should be higher if the activity is known to be inherently risky and attractive to risk seekers. But subjects may not be sensitive to such information. We shall manipulate this variable by descriptions of the defendant, e.g., as an "investment banker who likes to put money on long shots" or "a government employee nearing retirement." We shall check the manipulation by asking about the risk attitude of the defendant.

## 3.3 Positive incentive effects

The term "incentive" refers generally to both deterrence and positive incentives for desirable behavior, i.e., reinforcement in Skinner's sense. We shall carry out experiments on positive incentives.

Punishment, in all parts of the law that employ it, is effective only if it provides incentive not to cause harm. (We put aside the benefits of prison in preventing crimes while the criminal is incarcerated, and its purported benefit of rehabilitation.) Similarly, reward is effective as incentive, as in the case of giving patents (or prizes) for inventions.

Just as subjects may ignore probability in assessing penalties, they may ignore incentive effects more generally. We have found this result for punishment (Baron & Ritov, 1993), but we want to extend this research to reward as well as punishment.

### *3.3.1 Incentive in property*

Property rights are economically justified because they provide incentive for development and maintenance of property. Ideally, these rights have no justification when they cannot provide this incentive. We hypothesize, though, that people will think of property rights as inherent in the relation between people and what now counts as property.

We shall devise scenarios that manipulate both the relevance of incentive and the current status of property rights. This is possible because some scholars (e.g., DeSoto) have argued that further extension of property rights could ameliorate property in poor countries, e.g., giving slum dwellers rights to their land and housing. Others have argued that property rights are part of the solution to commons dilemmas such as overfishing. The latter example is particularly useful for our purposes, because we can present cases in which rights either provide incentive or not. For example, the right to fish in a certain small area of the ocean does not provide any incentive to conserve fish in that area, if the fish are constantly swimming in and out of it. But, if it is a closed bay, a property right will provide an incentive. We hypothesize that subjects' willingness to assign property rights will not depend on the relevance of incentive.

Intellectual property (IP) is another topic of great current interest. Many laws concerning IP - such as patent laws and, to some extent, copyright laws - provide incentive by granting to a creator a temporary monopoly. From an economic perspective (Shavell, 2003), this kind of law is a compromise between the benefits of providing incentive for creation and the costs of monopoly pricing.

Again, we shall manipulate both the incentive function of laws and their current status. (It is easy to find cases where IP laws are too strong, according to economic theory, and we may also be able to find cases where they are too weak.) For example, we shall ask about IP rights for works that will not be produced (writing out of print, patents that lead to no production). In these cases, the incentive function is reduced. We shall ask whether people favor weaker laws in these cases.

Another issue, much discussed, concerns cases in which creative ideas are used by others to create further ideas. IP restrictions in these cases have negative effects as well as positive ones. Hence, most (but not all) scientific research has no associated IP limitations on use (although it does have other rights, such as copyright of articles). Some think that computer software creation works much the same way and that too much software is protected from use by others. We shall ask whether people understand this conflict by presenting scenarios in which a type of computer program is either usable by others or not, and we shall also vary the effort involved in creating it. Arguably, when effort is high and potential use by others is low, the software should be strongly protected.

### *3.3.2 Probability effects with positive incentives*

We shall examine probability effects in property law. Arguably, protection of intellectual property (IP) should be greater when the probability of success of a creative effort is lower. This is the other side of the coin from the deterrent effect of punishment.

Invention or creation may be rewarded in several ways. In addition to copyrights, patents, and protection of trade secrets, Shavell (2003) suggests the possibility (tried at various times) of giving awards or prizes. The idea is to treat an invention as a public good, hence something to be supported by government rather than through (inefficient and distortionary) monopoly pricing. (Note that patents are also granted and enforced by government, so the main difference is who pays, not the extent of the bureaucracy.) We shall present subjects with this sort of hypothetical situation as a way of eliciting additional judgments, aside from asking them about duration of patents.

We shall ask for judgments about appropriate patent durations - or prizes - for creative work, such as developing a vaccine or new drug, or a computer program. We shall manipulate the expected benefit of the invention and the probability of success in the effort to make it. (For example, drug companies figure that a certain proportion of invented drugs will not be approved, and, even before that, a certain proportion of research efforts lead to no drugs.) We hypothesize that judgments of reward will depend heavily on the expected benefit (given that the invention is produced) and not enough on the probability of success of the effort to make it.

To insure that this is not simply a matter of the type of number being used, we can describe the expected success in terms of a probability, e.g., the probability of curing a given patient of a disease.

### 3.3.3 *Incentive effects in contracts*

Breaking a contract is like breaking a promise. Intuition might favor a strong penalty. Yet, economic analysis of the most common case - an incomplete contract between a buyer and seller - favors a more modest penalty based on the expected value of the contract to the buyer (typically the buyer's maximum willingness to pay), so that the seller has an incentive to *break* contracts when they should be broken, namely, when the seller's cost of keeping them exceeds the benefit to the buyer. A stronger penalty would be appropriate if the parties had a chance to negotiate a complete contract. Such a contract would specify that the seller is not obliged to provide the good being sold if the cost of producing it is too high (Shavell, 2003).

We hypothesize that many subjects would not be sensitive to this distinction when they judge the appropriateness of penalties for breach of contract, just as they are insensitive to incentive effects in tort penalties.

### 3.3.4 *Incentive in risk regulation*

The economic theory of risk regulation specifies an optimal amount of risk. "If you never miss a plane, you are spending too much time waiting in airports." Similarly, there is an optimal amount of pollution - a typical form of risk. The optimum depends on how the harm and the benefit vary with the amount of activity. In general, the harm is roughly linear with activity level, but the benefit is marginally declining. For example, if the production of a widget causes a fixed amount of pollution, some number of widgets produced might be very important to some people. But, as the number increases, the utility of the next widget will decrease to the point of being less than the disutility of the pollution.

In theory, the optimum can be achieved through a pollution tax, which makes the polluter pay all and only the cost of the external effects of the pollution. (The optimum can also be achieved in other ways, such as a subsidy for not polluting, or a sale of tradeable rights to pollute, sold after the optimum pollution level is calculated.)

It is considered immoral to put others at risk, however, so the idea of an optimal level of risk is non-intuitive. We hypothesize that subjects will be insensitive to the benefit of the activity in question when deciding between optimal methods and levels of risk regulation.

## 4. Resources

### ***Human subjects***

#### 4.1 Printed questionnaire studies

Questionnaires will be given, at the Hebrew university to undergraduate students in classrooms, or in experimental sessions. Subjects in the experimental sessions will be paid for participation.

#### 4.2 Web studies

This section is the general description for Jonathan Baron's web studies, which has been in use for several years.

**Subject recruitment:** The studies are all conducted on the World Wide Web. Subjects are paid by the study, usually \$3. The usual time to complete such a questionnaire is 15-20 minutes. Subjects who go too quickly are eliminated. (They are warned about taking the study seriously.)

The current instructions to subjects are in (<http://www.psych.upenn.edu/~baron/qs.html>). Because I have been doing these studies for a few years, I do not need to recruit. Links exist to my studies in many web pages maintained by others (and not solicited by me), such as those that keep lists of how to "make money by surfing the web." My studies are also discussed in web discussion lists, and in general I have a good reputation. You can find these by searching in Google for my name and "questionnaires." The instructions now specify that the studies are limited to people 18 or older.

My main problem has not been recruiting subjects but, rather, limiting them. I do not want to just pay less. I don't think that is fair. My studies are quite difficult and demanding. I also require serious answers (and I warn people of that). For example, I do not pay people who simply give the same answer to every question. Thus, recently I have adopted a different system, in which I have a "panel" of 2,000 registered subjects. I send them email when a study is ready - usually 500 at a time so that everyone has a reasonable chance of doing it within the few hours it is available. This seems to be working well. Subjects like it.

**Past findings:** My web page contains the full text of many papers that grew out of this research program, including published papers. There is too much to attempt to summarize here. See <http://www.sas.upenn.edu/~baron>.

**Potential risks:** None beyond the risks of daily life on the internet (computer crashes, misunderstanding, mail not getting through, etc.). Recently I have improved matters immensely by writing a new PHP interface for submission of data. I also added a fail-safe mechanism, which allows subjects to submit data by email if this fails, which has not happened in the last 1,000 or so submissions. The biggest risk for subjects is to complete a study and not be able to submit it, hence not get paid. I think have reduced this to zero.

**Consent procedures:** Clicking on the link to the questionnaire amounts to consent. The instructions in the main page are equivalent to a consent form. I have not used a formal consent form in the 29 years I've been at Penn, because I have argued (and still believe) that it is more likely to arouse anxiety than to allay it. But I do very much care about providing people with full information about what they are getting into (and I am constantly revising my page when I notice that I seem not to be getting through to everyone).

**Protection of subjects (confidentiality):** In addition to doing battle with the business office and Accounts Payable, to try to get the subjects paid in some sort of timely manner, my main concern is confidentiality. In order to get paid, the subjects must provide their name, address, and (if in the U.S.) social security number. They submit this on a secure registration form. I compile this information into a database, as described in <http://finzi.psych.upenn.edu/~baron>. I do keep two copies of this database, one on my home computer and one on my office computer ([finzi.psych.upenn.edu](http://finzi.psych.upenn.edu)), both in read-protected directories.

Data from individual studies contain only email addresses as identifying information. I use these to link up to my database in order to pay people. I also save *one* copy of the raw data - kept in a protected directory in [cattell.psych.upenn.edu](http://cattell.psych.upenn.edu) until subjects are paid, and then moved (not copied) to my home computer. I need to keep this in case questions arise about pay (as they often do). It is my most authoritative record of who has done what. Again, there is only one copy of this on a computer that is extremely secure.

After a study is done, I also make a reduced copy of the data for analysis purposes, with all identifying information removed. I make several backups on this, on several different computers (my home computer, my office computer, a laptop, and sometimes on other computers when the study is collaborative). But, again, all identifying information (including the email address) has been removed at this point.

**Benefits for subjects:** In addition to the pay, many subjects say they find my studies interesting, in the way that solving puzzles is interesting. Some subjects find them boring. I assume they do not come back for more. Occasionally, I get involved in lengthy e-mail correspondence with a subject. In one case, the subject in question applied to graduate school with the idea of working with me. (Ultimately, she went elsewhere, although she was admitted.)

**Risks-benefit ratio:** For the subjects who actually do participate, it seems that the benefits outweigh the risks. Even if they find the study boring, they find the small amount of money I offer (usually \$3 per study) to be adequate compensation. I rarely get complaints about my rate of pay.

##### **5. Short report of the results in our currently funded project: "Evaluation of public action: Attitudes and protected values"**

Decisions concerning public actions, just like other decisions, should normatively lead to the best consequences. However, we know of too many instances violating this principle. These violations can result from the direct effect of emotions or they may reflect a decision process that is non-consequential in essence. For example, the decision may be based on attitudes, values, or rules concerning the action, rather than its consequences. In our current research we have explored the role of these factors in decisions about public issues.

In recent years there has been growing interest in the different roles emotions play in the decision making process. While the affective system may have an important task in directing the individual's attention to potentially high priority events, its role is not entirely positive. Indeed, the direct emotional influence on the decision process (including feelings such as worry, fear, dread and anxiety) often result in behavioral responses that depart from what the individual would view as the best course of action (Slovic et al., 2002, Loewenstein et al., 2001). In particular, stimuli that evoke more immediate emotional response may take

precedence over less affect rich stimuli, even if the latter have more important consequences.

In our research we have explored the possibility that emotions cause biases by acting more directly on the judgment or decision. We were not concerned just with the imbalanced attention to emotion laden options or the insufficient notice of relatively unemotional ones, but rather in the direct effect of emotional response to the evaluation target. We also explored the relations between emotions and protected values, namely those values that people think of as absolute, not to be traded off for anything else.

The first set of experiments concerned the effects of emotion on judgments and decisions involving consequences. These experiments show that emotions, particularly bad emotions, are associated with non-consequentialist judgments and decisions. We included PVs as examples of such decisions, even though, technically, they could be normatively consistent with someone's utilities.

The second set of experiments concerned the determinants of emotions and their effects. We found that different kinds of response measures increase or decrease the involvement of emotion. In particular, when items are presented jointly, emotions are evoked by both items and may tend to conflict with each other, so that each emotion is weaker than it would be if an item were presented in isolation. Thus, emotions have less effect when items are presented jointly.

We have drafted a very long report of this work, which may end up as more than one paper. It is not quite ready for public view.

We have submitted another paper, which has been tentatively accepted by *Organizational Behavior and Human Decision Processes*, about omission bias and "normality." It is a response, in the form of a literature review and a new empirical study, to articles by Connolly and Reb and by Prentice and Kohler concerning alternative interpretations of our earlier work.

## 6 References

- Babcock, L., Gelfand M, Small D, Stayn H. (2003). The propensity to initiate negotiations: toward a broader understanding of negotiation behavior. Manuscript, Carnegie-Mellon University.
- Baron, J., & Hershey, J. C. (1988). Outcome bias in decision evaluation. *Journal of Personality and Social Psychology*, 54, 569-579.
- Baron, J., & Ritov, I. (1993). Intuitions about penalties and compensation in the context of tort law. *Journal of Risk and Uncertainty*, 7, 17-33.
- Birnbaum, M. H. (Ed.) (2000). *Psychological Experiments on the Internet*. New York: Academic Press.
- Birnbaum, M. H. (1999). Testing critical properties of decision making on the Internet. *Psychological Science*, 10, 399-407.
- Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, 37, 48-74.
- Brennan, G., & Lomasky, L. (1993). *Democracy and decision: The pure theory of electoral politics*. Cambridge: Cambridge University Press.
- Dudoit, S., & Ge, Y. (2003). Bioconductor R packages for multiple hypothesis testing: multtest. <http://www.bioconductor.org>.
- Eisenberg, T., Goerdts, B O., Rottman, D., and Wells, M. T. (1997). *Journal of Legal Studies*, 26, 623-661.
- Fox, C. R., & Birke, R. (2002). Forecasting trial outcomes: Lawyers assign higher probability to possibilities that are described in greater detail. *Law & Human Behavior*, 26, 159-173.
- Green, D. P. & Shapiro, I. (1994). *Pathologies of rational choice theory: A critique of applications in political science*. New Haven: Yale University Press.
- Loewenstein, G., Weber, E., Hsee, C., & Welch, N. (2001). Risk as feelings. *Psychological Bulletin*, 127, 267-286.
- McGraw, K. O., Tew, M. D., Williams, J. E. (2000). The integrity of Web-delivered experiments: Can you trust the data? *Psychological Science*, 11, 502-506.
- Polinsky, A. M. (1997). Are punitive damages really insignificant, predictable, and rational? A comment on Eisenberg et al. *Journal of Legal Studies*, 26, 663-667.
- Polinsky, A. M., & Shavell, S. (1998). Punitive damages: An economic analysis. *Harvard Law Review*, 111, 870-962.
- Rachlinski, J. J. (1998). A positive psychological theory of judging in hindsight. *University of Chicago Law Review*, 65, 571-625.
- Shavell, S. (2003). *Foundations of economic analysis of law*. Cambridge, MA: Harvard University Press.
- Slovic, P., Finucane, M., Peters, E., & MacGregor, D. G. (2002). The affect heuristic. In Gilovich, T., Griffin, D., and Kahneman, D. (Eds.). *Heuristics and Biases*. New York: Cambridge University Press.
- Sunstein, C. R., Schkade, D. & Kahneman, D. (2000). Do people want optimal deterrence? *Journal of Legal Studies*, 29, 237-254.
- Tykocinski, O. E. (2001). I never had a chance: Using hindsight tactics to mitigate disappointments. *Personality and Social Psychology Bulletin*, 27, 376-382.
- Ubel, P. A., Baron, J., & Asch, D. A. (2001). Preference for equity as a framing effect. *Medical Decision Making*, 21, 180-189.
- Westfall, P. H., & Young, S. S. (1993). *Resampling-based multiple testing: Examples and methods for p-value adjustment*. New York: John Wiley & Sons, 1993.