Experimental Measurement of Attitudes Regarding Cybercrime

Experimental Measurement of Attitudes Regarding Cybercrime*

James T. Graves
Carnegie Mellon University

Alessandro Acquisti
Carnegie Mellon University

Ross Anderson
University of Cambridge

May 22, 2014

Abstract

We conducted six between-subjects survey experiments to examine how judgments of cybercrime vary as a function of characteristics of the crime. The experiments presented vignettes that described a fictional cybercrime in which someone broke into an organization's network and downloaded data records. In each experiment, we manipulated the vignettes according to one dimension per experiment: type of data, scope, motivation, the organization's co-responsibility for the crime, consequences, and context. Participants were U.S. residents recruited via Amazon Mechanical Turk. We find that scope (the number of records downloaded) and the attacker's motivation had significant effects on the perceived seriousness of the crime. Participants also recommended harsher punishments when the monetary costs of the cybercrime were higher. Furthermore, participants considered cybercrimes committed by activists to be significantly less blameworthy, and deserving of significantly lighter sentences, than cybercrimes committed for profit—contrary to the position sometimes taken by U.S. prosecutors.

1 Introduction

The past few years have seen a growing interest in combining cybercrime economics and decision research to understand the causes and consequences of cybercrime. Much effort has been directed towards studying attacker behavior [8, 13, 17, 12], attacker psychology [20, 33, 26], and the psychology of victims, particularly with respect to the biases attackers exploit in their victims [40, 23].

To our knowledge, attitudes towards cybercrime have been relatively less explored. But how society—including victims and potential victims of cybercrime—views cybercrime is an important topic for study. Public perceptions of crime can affect how cybercrimes are defined, what punishments they carry, whether those punishments are believed to be fair, and how resources are allocated to enforcement [27].

Accurately assessing perceptions of cybercrime is especially important considering that cybercrimes may be punished differently from equivalent real-world crimes [16]. Consider, for example, the case of Matthew Keys, a journalist who gave a member of Anonymous the password to the Chicago Tribune's content management system. Anonymous used the access (which they had for about 30 minutes) to change a story online. Keys was charged with three felonies, which a U.S. Department of Justice press release claimed carried a combined maximum sentence of up to 25 years. The similar real-world crime of vandalism is punishable in California by a maximum of 3 years, if charged as a felony. Some have, in fact, argued that the punishments for cybercrime are not

*This work was partially funded by the Department of Homeland Security (DHS) Science and Technology Directorate, Cyber Security Division (DHS S&T/CSD) Broad Agency Announcement 11.02, the Government of Australia, and SPAWAR Systems Center Pacific via contract number N66001-13-C-0131.

harsh enough: a writer for Slate argued that cybercriminals should be given the death penalty.[1] On the other hand, cybercrimes for profit are sometimes punished less severely than the corresponding real-world offences. In fact, a case that started the discussions which led to the work described in this paper took place in March 2011, when a student at Greenwich University defrauded a professor at Oxford of £18,000—and got a sentence of 120 hours community service.[2] The perpetrator had been a mule in a phishing scam, so the fraud was online, and victim and attacker never met. Had it been face-to-face, sentencing guidelines would have set the starting point at 3 years.

Our investigation contributes to that debate by offering the first analysis (of which we are aware) of individuals' attitudes and perceptions of cybercrime. Our work extends the literature on general crime seriousness [41, 32] by exploring how attributes of a particular type of cybercrime affect perception of that crime. Our work is also inspired by the burgeoning field of experimental philosophy, which uses experimental techniques to measure attitudes about ethical questions.

We measure attitudes towards cybercrime using a series of six survey experiments. Each experiment randomly assigned participants to one of two or more conditions. Participants assigned to any condition answered a survey that presented a vignette description of a fictional cybercrime. The vignettes varied according to the condition a participant was assigned. The dimensions on which we focused were the type of data, scope, motivation, the organization's co-responsibility for the crime, consequences, and context. We find that scope (the number of records downloaded) and the attacker's motivation had significant effects on the perceived seriousness of the crime. Participants also recommended harsher punishments when the monetary costs of the cybercrime were higher. Perhaps most interestingly, our participants (who were U.S. residents) considered cybercrimes committed by activists to be significantly less blameworthy, and deserving of significantly lighter sentences, than cybercrimes committed for profit—contrary to the position sometimes taken by U.S. prosecutors.

2 Background

Our study uses experimental methods to measure attitudes about cybercrime. As such, our work is related to streams of research in three areas: (1) cybercrime economics, (2) perceptions of crime seriousness, and (3) experimental philosophy.

2.1 Cybercrime Economics

Cybercrime is a significant global problem, and much research has gone into understanding it.

Anderson et al. [1] developed a framework for measuring the total economic cost of cybercrime. They observe that a challenge in finding the total cost of cybercrime is defining "cybercrime." They discuss three types of cybercrime: (1) genuine cybercrime that is not possible without the internet, (2) traditional crimes that happen to occur over the Internet now instead of face-to-face, and (3) "transitional" cybercrime, which is a sort of hybrid in which the Internet has greatly changed how traditional attacks are now carried out. Anderson and his co-authors categorize the cost of cybercrime as consisting of criminal revenue, direct losses, indirect losses, and the cost of defense. Although the authors caution against simply adding up their numbers to obtain a "total cost of cybercrime," it is clear that the costs of payment card fraud alone run into the multiple billions of dollars.

[1] Steven E. Landsburg, Feed the worms who write worms to the worms: The economic logic of executing computer hackers. *Slate*, May 26, 2004.

[2] Paul Cheston, Criminology student helped defraud professor. *London Evening Standard*, Mar 28 2011

It is unsurprising, then, that much work has gone into trying to understand the economics of cybercrime. Some researchers have focused on the markets, studying particular black market sites such as SilkRoad [8], the marketplaces for specific goods like pharmaceuticals [25], or on crime in crowd-sourced labor markets [13].

Cyberattacker psychology has also been a subject of much research. Nykodym [26] looked to build profiles of insider cybercriminals. Rogers [33] studied the psychology and behavior of self-reported computer criminals. The study of cybercriminal psychology has even reached the point to have spawned at least one textbook on the subject [20].

Researchers have also studied victim behavior. Stajano and Wilson [40], for example, looked at the psychological reasons people fall for (online) scams. Stephen Lea in the UK produced an extensive report [23] on the psychology of scams for the UK Office of Fair Trading.

Others have focused on the economics of penalties for cybercrime. For example, Png, Wang, and Wang [30] studied how enforcement affects cybercrime. They found that enforcement mainly led to displacement: greater enforcement in the U.S. increased attacks coming from other countries. They also found a connection between the number of attacks and the U.S. unemployment rate.

2.2 Crime Seriousness

The literature on perceptions of crime seriousness dates back to the 1960s (for overviews, see [41, 32]).

Sellin and Wolfgang [37] used surveys of judges, police, and college students in Philadelphia to rank the seriousness of fifteen criminal acts. Several authors have reproduced this work while correcting potential methodological issues with the original study. Walker [43] confirmed Selin and Wolfgang's results with a sampling of the general population. She also used different assessment methods to make sure that the same subjects' ratings were consistent.

Rossi et al. [36] surveyed adults in Baltimore to obtain crime serious ratings for 140 offenses. Riedel [31] explored "what effect differing perceptions of circumstances of the crime have on seriousness judgments." Blumstein and Cohen [5] surveyed a random sample of residents of Allegheny County, Pennsylvania to obtain opinions of the proper sentence for different crimes. These judgments were compared to judgments of crime seriousness and actual sentencing practices. Cohen [9] used actual injury rates and jury awards to calculate the harmfulness of various crimes to their victims; he compared these results to public perceptions of crime severity.

Howe [18] and Warr [44] each explored the dimensions that comprise crime seriousness. Howe analyzed crime seriousness along twelve dimensions. Warr focused on two dimensions of seriousness: the harmfulness of the crime and its moral wrongfulness.

Several researchers have examined perceptions of white-collar crime in particular. Cullen et al. [11] appear to have been the first to do so, with a survey in of residents of a small Illinois town, in the style of Rossi [36]. Rosenmerkel [34] extended Warr's work [44] by exploring the harmfulness and wrongfulness as components of crime seriousness in the context of white-collar crime seriousness. They found that survey participants judged white-collar crimes based on the crimes' harmfulness and wrongfulness, with a heavier reliance on the former.

2.3 Experimental Philosophy

This work is also inspired by the field of experimental philosophy [21]. Experimental philosophy "uses methods normally associated with psychology to investigate questions normally associated with philosophy" [22]. Those methods often involve between-subjects experiments in which participants are presented with vignettes that are manipulated across conditions. For example, Gino et

al. [14] used between-subjects scenario designs to study the outcome bias in ethics. In another study, participants were asked to make moral evaluations about a scenario that described a chairman of a company making a decision about a product that, depending on the experimental condition, would hurt or help the environment [21].

A classic problem in experimental philosophy is the "trolley problem" in ethics [42]:

> Frank is a passenger on a trolley, whose driver has just shouted that the trolley's brakes have just failed, and who then died of the shock. On the track ahead are five people; the banks are so steep that they will not be able to get off the track in time. The track has a spur leading off to the right, and Frank can turn the trolley onto it. Unfortunately there is one person on the right-hand track. Frank can turn the trolley, killing the one; or he can refrain from turning the trolley, letting the five die.

Most people might say that they would choose to let the one die to save the five. Experimenters have tested the extent to which responses change with context, cultural biases, and other factors [4].

3 Methodology

3.1 Research Questions

Our primary research goal is to study perceptions of cybercrime. But cybercrime takes many forms and has many features. We designed six between-subject experiments in which we manipulated different relevant features of cybercrime, one at a time. Each experiment relied on the presentation of a vignette describing an intentional data breach of consumers' personal information.

We chose an intentional breach of data for a number of reasons. The data breach scenario is one that we believe is readily understandable by most people. It also lends itself to manipulation of the attributes in which we are interested (scope, context, motivation, etc.) while holding other attributes (mostly) constant. It is also quite common.[3]

For the purposes of this study, we chose to focus on six aspects of cybercrime likely to influence perceptions of the crime. We expected each aspect to be related to either the wrongfulness or harmfulness of the cybercrime:

- *Type of data.* We would expect the theft of sensitive data to be seen as more wrongful and potentially more harmful than the theft of less-sensitive data.
- *Scope.* The greater the scope of a cybercrime, the more harmful it is likely to be. With a data breach, the scope is (in part) the number of records downloaded. We would expect, therefore, to see perceptions of crime harmfulness and severity increase as the number of records downloaded in a data breach increases.
- *Motivation.* We expect that people consider motive when assessing crime seriousness. We would expect a cybercriminal with a profit motive to be perceived as worse than a political activist or someone doing it to learn how.
- *Consequences.* Crimes are often judged according to their consequences. We hypothesize that a cybercrime with more expensive consequences would be seen as more harmful, but not necessarily more wrongful, than cybercrimes that result in lower damages. The concept

[3] According to the Privacy Rights Clearinghouse Chronology of Data Breaches, over 600 million data records have been affected in over 4,000 data breaches since 2005. See http://www.privacyrights.org/data-breach.

of consequences includes not only the amount of losses, but to whom they occur. We would expect people to perceive cybercrimes to be worse when large losses are suffered by consumers compared to large losses by businesses.

- *Co-responsibility.* The organization from which records are downloaded may bear some of the responsibility for its disclosure and use. Our hypothesis is that an organization that did had not taken reasonable security measures when it was breached would be perceived as more co-responsible for the cybercrime than an organization that as diligent about security.

- *Context.* We also wondered if context mattered: is downloading data from a bank worse than downloading it from a government agency or a non-profit?

One of the challenges in designing the study was to manipulate only one attribute at a time. A change in scope, motivation, or type of data, for example, can easily also affect implied consequences: when more records are downloaded, more people might be affected. Or when an attacker is motivated by greed instead of activism, one might assume that he is more likely to sell the data and that the data is therefore more likely to be misused. Or, again, credit card data can be used to make fraudulent purchases; health data may be sensitive, but the harm from its disclosure is more ambiguous and less financial. We were therefore careful to choose vignette language that minimized the possibility that a manipulation of one variable would "spill over" into an effect on consequences, which we believed might dominate other manipulations. At the same time, vignettes had to be believable. We addressed these issues by specifying consequences, where possible, and by using a vignette in which the perpetrator of the data breach did not actually release the data he downloaded. This design also had the desirable side effect of limiting extreme "ceiling" effects in the responses to our questions. Because the consequences were held to a minimum, the answers in each vignette were better distributed across the range than they otherwise might have been.

3.2 Design

Our design consisted in six between-subjects online survey experiments. Within each experiment, participants were randomly assigned to one of the conditions in that experiment. Depending on the experiment, the number of conditions ranged from two to five. The six experiments manipulated:

1. *Type of breached data.*
2. *Scope of breached data.*
3. *Motivation underlying the data breach.*
4. *Consequences of the the data breach.*
5. *Co-responsibility for the data breach.*
6. *Context of the data breach.*

All experiments (and their conditions) followed the same structure. Participants who passed a screening process were presented with an online survey; the survey asked them to read a vignette similar to the following:

> On June 3, 2013, while browsing the Internet, Tom Smith discovered a security flaw in the Acme Insurance Company's website. He used that flaw to gain access to Acme's internal network and download 100,000 records from Acme's customer database. Each record consisted of a customer's full name, phone number, and address. Tom did not use or release the information. Acme's customers suffered no harm.

Each experiment modified or extended this vignette with a particular manipulation. In the "Scope" experiment, for example, the survey presented the number of records downloaded as 10, 100, 1,000, 10,000, or 1,000,000 records depending on condition. In the "Context" experiment, the organization from which Tom Smith downloaded the data was described as a bank, a non-profit organization, or a government agency. In the "Motivation" experiment, the vignette included text explaining why Tom Smith was looking for security flaws—he was trying to make money, was a student looking to learn about computer security, or was an activist looking for evidence of corporate corruption. And so on with the other experiments.

After they read the vignette, participants were presented with a series of multiple-choice questions intended to test their recall of (and attention to) the details of that scenario. Each experiment included questions to test participants' recollection of the vignette's data type, context, and scope. If the manipulation was not covered by one of these three questions, an additional question was added to check recall of the manipulation. After each memory-check question, each participant was presented with a page indicating whether his or her answer was correct. In either case, the correct answer was repeated. This reminder was designed to further reinforce the participant's awareness of the details of the scenario.

Following the memory check questions, the survey collected the variables of interest. Participants were asked to answer a series of questions on a 1–7 Likert scale. We selected the first three questions in accordance with previous research on the factors of crime seriousness [44]. The survey presented the questions in random order. Variable names follow each question:

- "How wrongful were Tom Smith's actions?" (`how_wrongful`)
- "How serious was the crime Tom Smith committed?" (`how_serious`)
- "How harshly should Tom Smith be punished?" (`how_harshly`)
- "How harmful were Tom Smith's actions?' (`how_harmful`)
- "How responsible was the Acme Insurance Company for the crime?" (`how_responsible`)[4]
- "How clever was Mr. Tom Smith?" (`how_clever`)
- "How sensitive was the data that Tom Smith downloaded?" (`how_sensitive`)

The survey also asked participants to recommend a specific punishment for the crime, assuming Tom Smith were convicted of the crime. The question was multiple choice, with 11 options ranging from no punishment at all on the low end, to probation, to a sentence of 0–30 days, all the way up to a sentence of life in prison on the high end, with intermediate sentence lengths in between.

In the Motivation, Consequences, Co-Responsibility, and Context experiments, the survey followed the specific-punishment question with a question about the *potential* consequences of Tom Smith's actions. This question was intended to help determine whether participants might be judging scenarios by potential consequences instead of the actual consequences that were described in the scenarios. The added question also made an additional attention check possible: participants who rated the potential consequences as lower than the actual consequences may not have been paying enough attention to the questions. Responses that did so were removed from the response set.

The next section of the survey included several questions intended to measure participants' attitudes and experiences about data protection and personal privacy. We used the 15-question Concern for Information Privacy (CFIP) scale [39]. We also asked how often participants had suffered identity theft (if at all), how often they provide fake information when registering for web sites, and how much they have heard or read about "use and potential misuse of information

[4]In the Context experiment, the"Acme Insurance Company" was replaced by "ACR."

collected from the Internet" in the past year. The survey instrument concluded with demographic questions and a few open-ended questions.

Details of the survey instrument appear in Appendix B.

We ran ordered probit regressions on each variable of interest. Regressions included controls for demographics, memory check correctness, and privacy attitudes. We treated the demographic variables `gender`, `us_birth`, `age_category`, `education`, `occupation`, `work_situation`, and the memory check variables as categorical variables. We treated as continuous variables the extent to which participants had been affected by cybercrime or privacy invasions (`freq_aff_by_cybercrime`), and the extents to which they use fake personal information (`fake_personal_info`) and are aware of media coverage of data misuse (`media_awareness`).

4 Results

For each experiment, we used Amazon Mechanical Turk (MTurk) to recruit participants 18 years of age or older who lived in the United States and had at least a 95% approval rating on MTurk. The demographics and data quality of MTurk experiments has been extensively studied in multiple experimental contexts [6, 10, 15, 24, 28, 35, 38]. Several studies have shown that recruitment for online studies through MTurk can lead to more representative samples and better data quality than studies using other "convenience" samples such as university students [2, 3, 7]. In a recent article, Peer et al. [29] found that reputation alone is often enough to ensure sufficient data quality in MTurk studies.

Each study described the task as taking "a short survey on crime." We screened potential participants to exclude anyone who had already been in any of the other experiments (or the pilot studies that preceded them) from participating in subsequent experiments in this series. We also filtered out any responses with duplicated IP addresses or MTurk IDs,[5] demographic data that contradicted the eligibility requirements on MTurk (e.g., responses claiming that the participant was under 18 years old or resided outside the U.S.), or contradictory answers (as further discussed below). The total number of responses across all experiments was 2,440 after elimination of duplicate and unreliable responses.

4.1 Experiment 1: Type of Data (Data Sensitivity)

This experiment manipulated the type of data affected in the breach. The data was listed as either:

- name, phone number, and address (the "Low" or "Directory" condition), or
- name, health history, medical diagnoses, and prescription records (the "High" or "Medical" condition).

We recruited 250 participants from MTurk in October 2013. After eliminating duplicate, incomplete, and unreliable responses, 239 responses remained for analysis (median age category: 25–34; female: 51%; no statistically significant differences across conditions in terms of age, gender, education, occupation, or work situation). Figure 1 shows the distribution of responses to the seven main Likert-type questions. Table 2 in Appendix A details regression results.

As expected, participants rated names, health histories, medical diagnoses, and prescription records as more sensitive than names, phone numbers, and addresses ($p < 0.001$). The effect is strong as well as significant: 72% of participants in the medical-data condition rated the data as 7 ("Extremely sensitive") or 6 compared to 34% of those in the directory-data condition.

[5] This could happen if a participant attempted to restart or retake the survey within the same survey round.

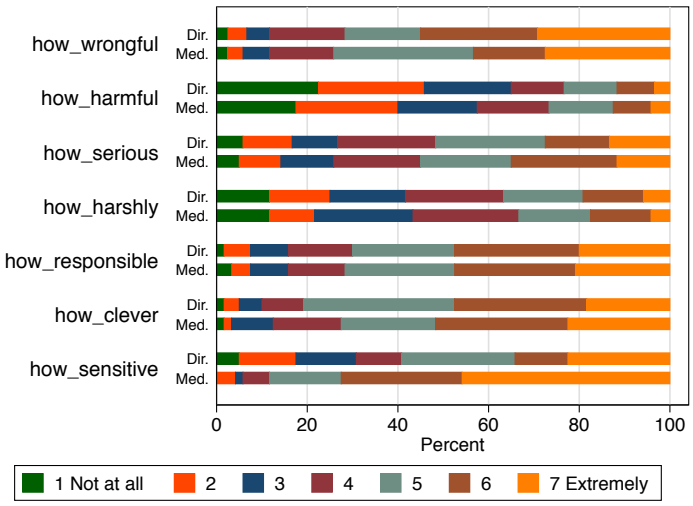

Figure 1: Responses to main Likert questions in the Type of Data experiment by condition

Perceived crime severity, however, did not differ between conditions with statistical significance. Answers to "How sensitive was the data?" and "How serious was the crime?" are strongly correlated ($p < 0.001$, χ^2), but the difference in perceptions of data sensitivity by condition did not translate to a statistically significant difference in perceptions of crime severity.

As would be expected, participants who are more concerned about privacy, as reflected in CFIP scores, rated the vignettes as more wrongful ($p < 0.001$), more serious ($p < 0.01$), and, at a one-sided significance of $p < 0.05$, more harmful. They also said that the crime should be punished more harshly ($p < 0.01$), and held Acme more responsible for the breach ($p < 0.05$). Women rated the crime more wrongful ($p < 0.01$), harmful ($p < 0.01$), and sensitive ($p < 0.05$) than did men. They also held Acme more responsible for the breach ($p < 0.05$).

Memory or attention to the scenarios had significant effects on some ratings, but a regression model that included interaction effects between these variables and condition showed no significant effect of answering the memory check question incorrectly across conditions.

4.2 Experiment 2: Scope

To experiment the effects of a crime's scope on perceptions, we manipulated the number of records the vignette said Tom Smith downloaded: 10, 100, 1,000, 10,000, or 1,000,000. All vignettes in this experiment described the records Tom downloaded as each containing "a customer's full name, phone number, address, date of birth, and social security number."

We recruited 625 participants from MTurk in November 2013. After eliminating duplicate, incomplete, and unreliable responses, 583 responses remained for analysis (median age category: 25–34; female: 41%; no statistically significant differences across conditions in terms of age, gender, education, occupation, or work situation). Figure 2 shows the distribution of responses to the seven main Likert questions. Table 3 in Appendix A details regression results.

Because the manipulated variable in this experiment is numerical, we included the condition in the model as a continuous variable representing the (base 10) logarithm of the number of records. The number of records had a statistically significant effect in the expected direction on all the Likert-type question responses. Note, however, that this may be due in part to the large sample

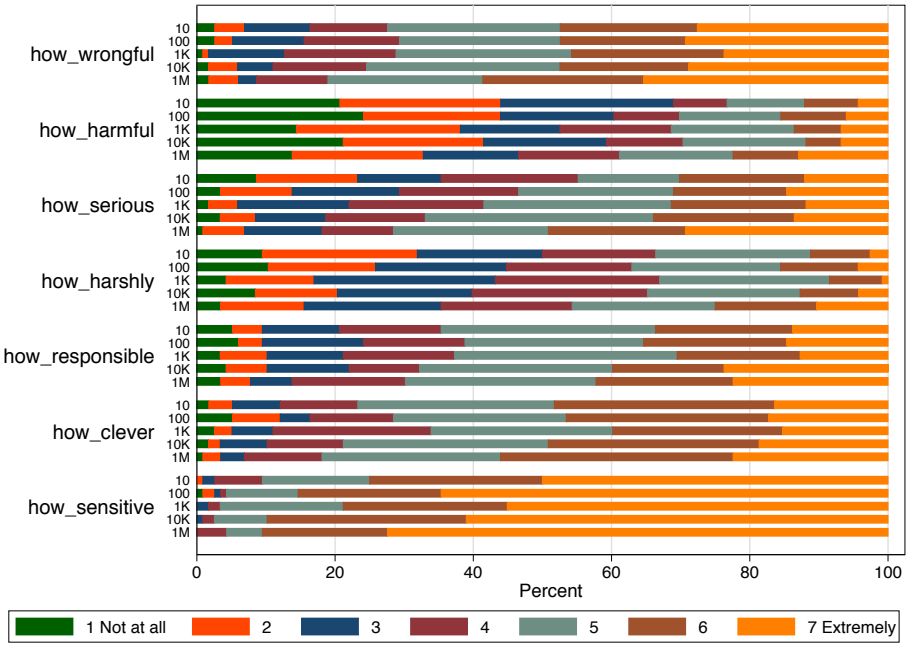

Figure 2: Responses to main Likert questions in the Scope experiment by condition

size compared to other experiments. Although the number of participants per condition is about the same as in other experiments, the total number makes it more likely that small-magnitude results such as those seen for Acme's co-responsibility for the breach ($\hat{\beta} = 0.064$, se $= 0.026$, $p < 0.05$) and Tom's cleverness ($\hat{\beta} = 0.058$, se $= 0.025$, $p < 0.05$) will be statistically significant.

Some of these results are surprising. One would expect perceptions of harmfulness and seriousness to increase with the number of records, and with them ratings of how harshly the crime should be punished. However, participants rated the data as more sensitive when more records were affected. The magnitude of that effect ($\hat{\beta} = 0.135$) is larger than that for any of the seven Likert questions, except for seriousness ($\hat{\beta} = 0.159$). Interpreting this result is challenging without additional information, but two possible explanations present themselves. First, the survey experiment may not have done an adequate job of asking about the sensitivity of the type of data downloaded as opposed to the sensitivity of the entire set of actual data records downloaded. Second, people may conflate sensitivity and the potential for harm from the data.

As in Experiment 1, the CFIP score is a strong correlate of ratings on all seven main Likert questions. Women's responses did not significantly differ from men's except for finding the data to be more sensitive ($p < 0.05$).

We note one curious anomaly that may be worthy of further investigation in future studies. In an initial pilot study we had previously completed, subjects considered it more wrongful to steal 100 medical records than to steal 10,000. We conjectured that this result might be a small-group effect, or the fact that "if I steal 100 medical records, I'm going to read them," as one interlocutor put it. There is a faint echo of that effect still visible in Figure 2: for example, taking 100 records is seen as "extremely" wrongful, serious, sensitive, and deserving of harsh punishment by more subjects than taking 1,000 records. However, this result is not significant, at least given the questions we asked in the main survey and the statistical tests we applied.

4.3 Experiment 3: Motivation

The motivation experiment manipulated the presentation of Tom Smith's reason for downloading the records. In this experiment, the vignette text explained that Tom was looking for security flaws:

> Tom Smith is a computer programmer who looks for security flaws on the Internet. He does this because he wants to `$motivation`. On June 3, 2013, Tom Smith found a security flaw in the Acme Insurance Company's website. He used that flaw to gain access to Acme's internal network and download 100,000 records from Acme's customer database. Each record consisted of a customer's full name, user ID, and password. Tom did not release the details of the flaw, and he did not use or release the records he downloaded. Acme's customers suffered no harm.

Tom's motivation was presented as wanting to:

- "learn about Internet security by looking for real flaws online and testing them" (the "Student" condition),
- "seek evidence of corporate corruption by looking for information inside corporate networks" (the "Activist" condition), or
- "make money by finding product designs, customer lists, and other trade secrets and selling them to the highest bidder" (the "Profiteer" condition).

As shown above, the vignettes described the data Tom downloaded as customers' full names, user IDs, and passwords. We also added a question asking how harmful the *potential* consequences might have been. As discussed in Section 3.2, this question was intended to help determine whether participants might be judging scenarios by potential consequences instead of the actual consequences described in the scenarios; the question also made an additional validation check possible.

We recruited 395 participants from MTurk in November 2013. After eliminating duplicate, incomplete, and unreliable responses, 361 responses remained for analysis (median age category: 25–34; female: 49%; no statistically significant differences across conditions in terms of age, gender, or education; occupation differed at $p < 0.05$ and work situation at one-sided $p < 0.05$, which we account for in the regressions). Figure 3 shows the distribution of responses to the seven main Likert questions. Tables 4 and 5 in Appendix A detail regression results.

Participants judged the profiteer more harshly than they did the student or activist. Participants rated the profiteer's crime as more wrongful ($p < 0.001$), harmful ($p < 0.05$), and serious ($p < 0.001$) than the student's, and said that it should be punished more harshly ($p < 0.001$). The difference between the profiteer and activist was only slightly less pronounced, with strongly significant results for both wrongfulness ($p < 0.001$) and seriousness ($p < 0.001$), and with one-sided significance for harmfulness ($p < 0.05$). The profiteer also received harsher judgments, compared with the activist, of how harshly he should be punished ($p < 0.01$).

Differences in perceptions of the student and activist are less pronounced. Although participants said that the activist should be punished more harshly than the student ($p < 0.05$), perceptions of wrongfulness, harmfulness, and seriousness were statistically indistinguishable.

CFIP is, again, a strong correlate of responses to the main Likert questions. Women held Acme more partially responsible for the crime than did men.

4.4 Experiment 4: Consequences

The consequences experiment manipulated the extent of damage from Tom Smith's actions. It also manipulated who bore the brunt of the damages. We added text to each vignette describing the

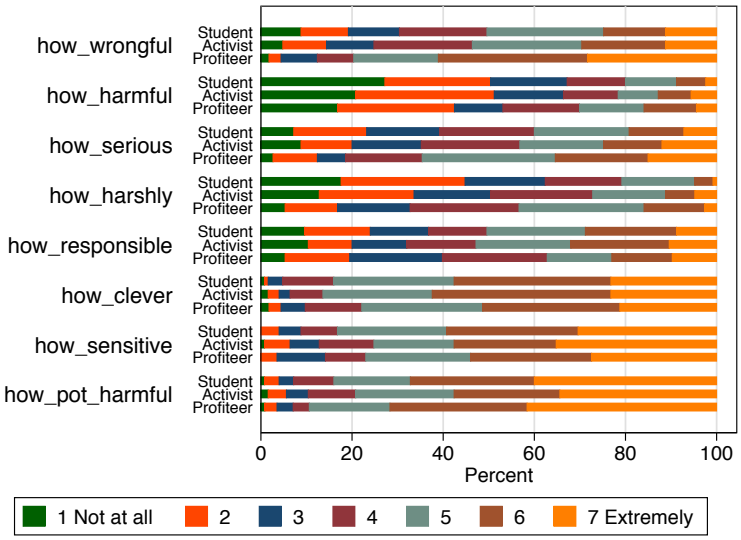

Figure 3: Responses to main Likert questions in the Motivation experiment by condition

harm that resulted from the crime. The text was one of the following:

- "Acme's engineers spent a few hours securing its servers at a cost of about $1000" (the "Low" condition),
- "Acme's engineers spent a few hours securing its servers at a cost of about $1000. Acme also spent $5 million to repair database damage" (the "Acme High" condition), or
- "Acme's engineers spent a few hours securing its servers at a cost of about $1000. Acme's customers spent a combined $5 million to protect themselves from identity theft" (the "Customers High" condition).

The vignettes said that the records "consisted of a customer's full name, address, and social security number."

We recruited 511 participants from MTurk in December 2013. After eliminating duplicate, incomplete, and unreliable responses, 479 responses remained for analysis (median age category: 25–34; female: 52%; no statistically significant differences across conditions in terms of age, gender, education, occupation, or work situation). Figure 4 shows the distribution of responses to the seven main Likert questions. Tables 6 and 7 in Appendix A detail regression results.

The manipulation had the expected effect on perceptions of harmfulness. The conditions in which either Acme ($p < 0.001$) or its customers ($p < 0.01$) spent $5 million received higher ratings of harmfulness than the condition in which the only cost was $1000 to secure servers. Participants also said that each of these two cases should be punished more harshly than the Low condition (Acme: $p < 0.01$, Customers: $p < 0.05$). Although the crimes involving $5 million loss were perceived to be more harmful than the Low condition, they were not perceived as more wrongful or serious with statistical significance (although the coefficients are in the expected direction).

Whether Acme or its customers bore the costs made little difference. Not only were the responses to the main Likert questions not statistically significant between the Acme High and Customer High conditions, the harmfulness of each condition was virtually the same ($\hat{\beta} = 0.03$, se $= 0.122$). This is somewhat surprising. We had expected that participants would empathize with customers over companies, and that empathy would mean rating damage to customers are more harmful than the

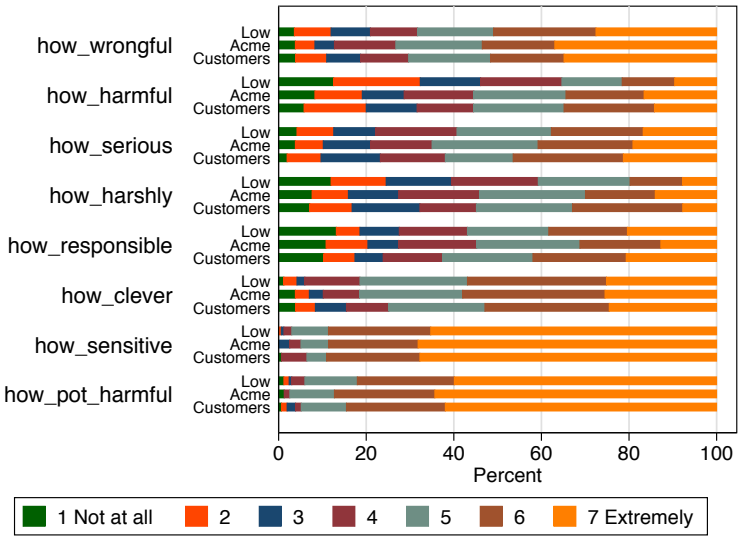

Figure 4: Responses to main Likert questions in the Consequences experiment by condition

same amount of damage to Acme, but this does not seem to have been the case. It could also be that people are more sympathetic to customers than companies, as we would expect, but that the two conditions are not as similar as we had hoped. $5 million in costs to a single company is not the same as $5 million in costs spread among 100,000 people.

CFIP is, again, a strong correlate of ratings of the main Likert questions. Women rated the potential harmfulness more severely than men did ($p < 0.05$).

4.5 Experiment 5: Acme's Co-Responsibility for the Crime

This experiment manipulated the steps Acme had taken to secure its own network. A sentence was added to the vignette depending on condition:

- "Acme had patched its server operating systems with the latest security updates" (the "Patched" condition), or
- "Acme had not patched its server operating systems with the latest security updates" (the "Not patched" condition).

This vignette described the data affected as names, addresses, and social security numbers.

We recruited 302 participants from MTurk in December 2013. After eliminating duplicate, incomplete, and unreliable responses, 276 responses remained for analysis (median age category: 25–34; female: 51%; no statistically significant differences across conditions in terms of age, gender, education, occupation, or work situation). Figure 5 shows the distribution of responses to the seven main Likert questions. Tables 8 in Appendix A details regression results.

The manipulation of whether Acme patched its servers has the expected effect on perceptions of the company's partial responsibility for the crime. Participants found Acme to share more responsibility for the crime when it had not patched its servers ($p < 0.01$). Participants did not find the crime significantly more wrongful, harmful, or serious when the company had not patched its servers, suggesting that they distinguish between a crime's seriousness of a crime and its causes.

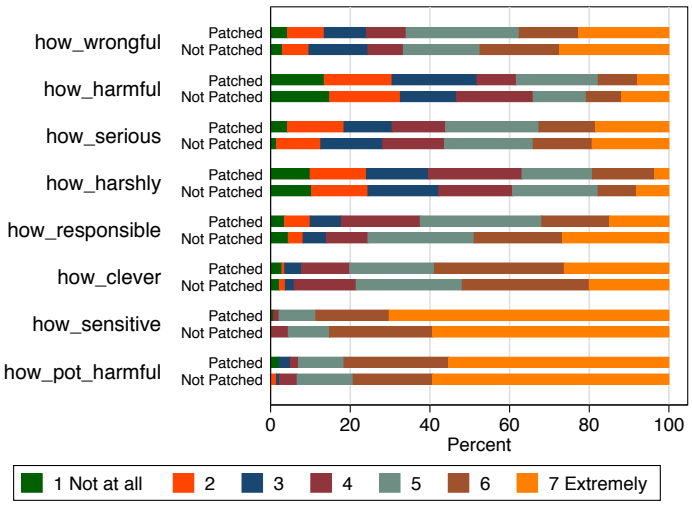

Figure 5: Responses to main Likert questions in the Co-Responsibility experiment by condition

In a surprising result, participants also rated the data as less sensitive when Acme had not patched its servers. Presumably some subjects take the view that the sensitivity of data best can be measured by the efforts taken to protect it.

4.6 Experiment 6: Context

The context experiment manipulated the vignette's description of the type of organization Tom Smith broke into: a bank, government agency, or non-profit. The vignettes listed the type of data downloaded as each customer's full name, address, and social security number. To try to make the perceived sizes of each type of organization the same, we included in each vignette that the organization had an "operating budget of $100 million per year." We also added a question about the size of the organization to use as a control: "How small or large of an organization do you consider ACR to be?"

We recruited 552 participants from MTurk in December 2013. After eliminating duplicate, incomplete, and unreliable responses, 502 responses remained for analysis (median age category: 25–34; female: 45%; no statistically significant differences across conditions in terms of age, gender, education, occupation; work situation differed between conditions at $p < 0.05$, which we account for in the regressions). Figure 6 shows the distribution of responses to the seven main Likert questions. Tables 9 and 10 in Appendix A details regression results.

The context manipulation showed no two-sided statistically significant effects on any of the main Likert questions, except for how partially responsible the breached organization was. Participants judged the non-profit to be less responsible for the breach than they did the bank ($p < 0.01$) or the government agency ($p < 0.001$). Participants did rate the non-profit vignette as less serious than either the government or bank scenario with one-sided $p < 0.05$.

As with the other experiments, CFIP was a strong correlate for crime seriousness, harmfulness, and wrongfulness. Organization size was also positively correlated with how potentially harmful the crime was ($p < 0.05$), how sensitive the data was ($p < 0.01$), and how clever the attacker was ($p < 0.01$).

Although the manipulations did not show strong significant effects on crime seriousness, this

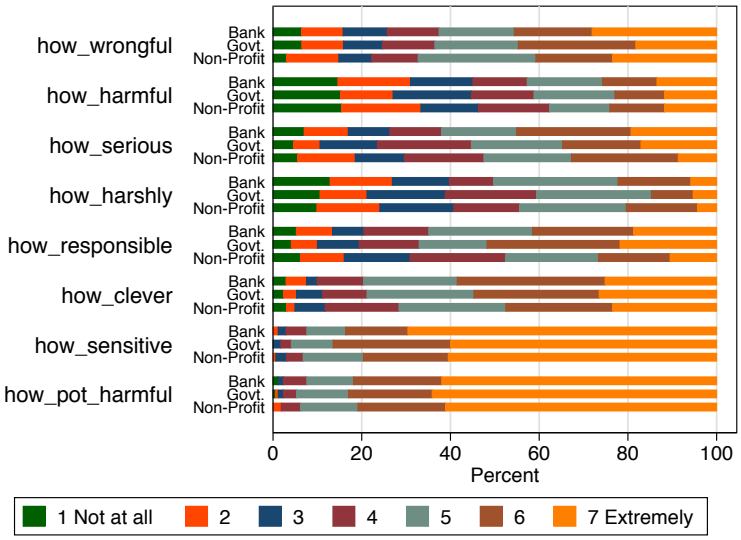

Figure 6: Responses to main Likert questions in the Motivation experiment by condition

result and the results of the experiment on co-responsibility suggest that the factors affecting perceptions of organizational responsibility to protect from breach are worthy of further consideration.

5 Discussion

Table 1 summarizes the statistically significant results of the regressions in all experiments.

Table 1: Summary of statistically significant regression results

Experiment & Conditions / How:	Wrongful	Harmful	Serious	Harshly	Pot. Harm.	Sensitive	Respons.	Clever
Type of Data: High v. Low					—	0.971***		
Scope: log(Records)	0.069**	0.078**	0.159***	0.106***	—	0.135***	0.064*	0.058*
Motiv.: Profiteer v. Student	0.877***	0.323*	0.593***	0.791***				
Motiv.: Profiteer v. Activist	0.793***		0.515***	0.485**				
Motiv.: Student v. Activist				−0.306*				
Conseq.: Acme v. Low		0.408***		0.341**				
Conseq.: Customers v. Low		0.377**		0.246*				
Conseq.: Customers v. Acme						0.252*		
Co-Resp.: Patched v. Not						0.364*	−0.420**	
Context: Gov't v. Bank								
Context: Bank v. Non-Profit:							0.359**	
Context: Gov't v. Non-Profit:							0.513***	

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table lists statistically significant results from ordered probit regressions in all experiments. "Pot. Harm" is marked off for the Type of Data and Scope experiments because that question was not asked in those experiments.

In various cases, the manipulations produced the expected effects. Changing the data from directory information to health information increased perceived data sensitivity. Increasing the number of records generally increased how wrongful, harmful, and serious the crime was perceived

to have been. Cybercrime committed by someone with a profit motive was rated as more wrongful than the same crime when committed by a person motivated by activism or a desire to learn. Respondents perceived an organization that had patched its servers to have been less responsible for the crime than an organization that had not patched its servers. Vignettes with more expensive consequences were judged to be more harmful. Downloading data from banks and government agencies was rated as more serious than downloading data from a non-profit.

The results support interpretations of cybercrime seriousness as composed at least partly of wrongfulness and harmfulness. Cybercrime vignettes that were rated as more wrongful were rated, with high statistical significance, as more serious, as were vignettes that were rated as more harmful. Data sensitivity, however, does not appear to be a major component of perceived cybercrime seriousness. Despite the data sensitivity in Experiment 1 having strongest effect of any manipulation, the perceived harmfulness, wrongfulness, and seriousness of the crime was not statistically significant across conditions.

Furthermore, Table 1 suggests that people distinguish between features of cybercrimes. Perceptions that data taken in a crime was more sensitive did not necessarily result in statistically significant perceptions that the crime was more serious or that the crime should be punished more harshly. As noted in the Introduction, one of the more interesting results is the comparative reaction of our participants (all U.S. residents) to cybercrimes committed by activists versus cybercrimes committed for profit. The former were considered significantly less blameworthy, and deserving of significantly lighter sentences, contrary to the position sometimes taken by U.S. prosecutors.

5.1 Future Work

Our analysis so far has focused on the effects of independent manipulations of single factors. In future work, we plan to use factorial vignette survey methodology [19] to explore the interactions between these factors.

The experiments we have discussed are all based on vignettes describing a data breach. But there are many types of cybercrimes, including payment card fraud, scamming, online banking fraud, phishing, and viruses. A natural extension of our work is to compare different types of cybercrime. In addition, we intend to study how cybercrimes are perceived in comparison with similar real-world crimes.

The surprising appearance of data sensitivity among statistically significant results of other manipulations suggests that perceptions of data sensitivity might be another area for future research. We also intend to study in more depth perceptions of fault on the part of breached organizations.

6 Conclusion

We described the results of six between-subjects survey experiments designed to examine how judgments of cybercrime vary as a function of characteristics of the crime. Those judgments matched in various cases our hypotheses: participants recommend harsher sentences when cybercrimes involve more sensitive data, cost more, or are motivated by profit. This is contrary to the current practice of prosecutors in many countries, including the USA and the UK, who are more vigorous in pursuing offenders and seeking long sentences where the offenders are students or activists rather than profiteers, and where the targets are government websites rather than those of commercial companies or nonprofits. One possible use of findings such as ours may be to help nudge sentencing practice towards what voters expect.

References

[1] Ross Anderson, Chris Barton, Rainer Böhme, Richard Clayton, Michael van Eeten, Michael Levi, Tyler Moore, and Stefan Savage. Measuring the cost of cybercrime. In *Workshop on the Economics of Information Security (WEIS)*, 2012.

[2] Tara S. Behrend, David J. Sharek, Adam W. Meade, and Eric N. Wiebe. The viability of crowdsourcing for survey research. *Behavior Research Methods (Online)*, 43(3):800–13, September 2011.

[3] Adam J. Berinsky, Gregory A. Huber, and Gabriel S. Lenz. Evaluating online labor markets for experimental research: Amazon.com's mechanical turk. *Political Analysis*, 20(3):351–368, 2012.

[4] April Bleske-Rechek, Lyndsay A. Nelson, Jonathan P. Baker, Mark W. Remiker, and Sara J. Brandt. Evolution and the trolly problem: People save five over one unless the one is young, genetically related, or a romantic partner. *Journal of Social, Evolutionary, and Cultural Psychology*, 4(3):115–127, 2010.

[5] Alfred Blumstein and Jacqueline Cohen. Sentencing of convicted offenders: An analysis of the public's view. *Law & Society Review*, 14(2):223–261, January 1980.

[6] Michael Buhrmester, Tracy Kwang, and Samuel D. Gosling. Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science*, 6(1):3–5, January 2011.

[7] Krista Casler, Lydia Bickel, and Elizabeth Hackett. Separate but equal? A comparison of participants and data gathered via Amazon's MTurk, social media, and face-to-face behavioral testing. *Computers in Human Behavior*, 29(6):2156–2160, 2013.

[8] Nicolas Christin. Traveling the silk road: A measurement analysis of a large anonymous online marketplace. In *Proceedings of the 22nd International World Wide Web Conference (WWW'13)*, pages 213–224, May 2013.

[9] Mark A. Cohen. Some new evidence on the seriousness of crime. *Criminology*, 26(2):343–353, 1988.

[10] Matthew J. C. Crump, John V. McDonnell, and Todd M. Gureckis. Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, 8(3):e57410, March 2013.

[11] Francis T. Cullen, Bruce G. Link, and Craig W. Polanzi. The seriousness of crime revisited: Have attitudes toward white-collar crime changed? *Criminology*, 20(1):83–102, 1982.

[12] Dinei Florencio and Cormac Herley. Where do all the attacks go? In *Workshop on the Economics of Information Security (WEIS)*, 2011.

[13] Vaibhav Garg, Chris Kanich, and L. Jean Camp. Analysis of ecrime in crowd-sourced labor markets: Mechanical Turk vs. Freelancer. In *Workshop on the Economics of Information Security (WEIS)*, 2012.

[14] Francesca Gino, Don A. Moore, and Max H. Bazerman. No harm, no foul: The outcome bias in ethical judgments. Working Paper 08-080, Harvard Business School, February 2008.

[15] Joseph K. Goodman, Cynthia E. Cryder, and Amar Cheema. Data collection in a flat world: The strengths and weaknesses of mechanical turk samples. *Journal of Behavioral Decision Making*, 26(3):213–224, 2013.

[16] Jennifer S. Granick. Faking it: Calculating loss in computer crime sentencing. *I/S: A Journal of Law and Policy for the Information Society*, 2:207–229, 2006.

[17] Chris Grier, Lucas Ballard, Juan Caballero, Neha Chachra, Christian J. Dietrich, Kirill Levchenko, Panayiotis Mavrommatis, Damon McCoy, Antonio Nappa, Andreas Pitsillidis, Niels Provos, M. Zubair Rafique, Moheeb Abu Rajab, Christian Rossow, Kurt Thomas, Vern Paxson, Stefan Savage, and Geoffrey M. Voelker. Manufacturing compromise: The emergence of exploit-as-a-service. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, CCS '12, pages 821–832, New York, NY, USA, 2012. ACM.

[18] Dmund Howe. Dimensional structure of judgments of crimes. *Journal of Applied Social Psychology*, 18(16):1371–1391, 1988.

[19] Guillermina Jasso. Factorial survey methods for studying beliefs and judgments. *Sociological Methods and Research*, 34(3):334–423, 2006.

[20] Grinne Kirwan and Andrew Power. *Cybercrime: The Psychology of Online Offenders*. Cambridge University Press, New York, NY, USA, 2013.

[21] Joshua Knobe. Intentional action and side effects in ordinary language. *Analysis*, 63(3):190–194, 2003.

[22] Joshua Knobe, Wesley Buckwalter, Shaun Nichols, Philip Robbins, Hagop Sarkissian, and Tamler Sommers. Experimental philosophy. *Annual Review of Psychology*, 63(1):81–99, 2012.

[23] Stephen Lea and Peter Fischer. The psychology of scams: Provoking and committing errors of judgement. Technical Report OFT1070, Office of Fair Trading, 2009.

[24] Winter Mason and Siddharth Suri. Conducting behavioral research on Amazon's Mechanical Turk. *Behavior research methods*, 44(1):1–23, 2012.

[25] Damon McCoy, Andreas Pitsillidis, Grant Jordan, Nicholas Weaver, Christian Kreibich, Brian Krebs, Geoffrey M. Voelker, Stefan Savage, and Kirill Levchenko. Pharmaleaks: Understanding the business of online pharmaceutical affiliate programs. In *Proceedings of the USENIX Security Symposium*, August 2012.

[26] Nick Nykodym, Robert Taylor, and Julia Vilela. Criminal profiling and insider cyber crime. *Computer Law & Security Review*, 21(5):408 – 414, 2005.

[27] Michael O'Connell and Anthony Whelan. Taking wrongs seriously: Public perceptions of crime seriousness. *British Journal of Criminology*, 36(2):299–318, 03 1996.

[28] Gabriele Paolacci, Jesse Chandler, and Panagiotis G. Ipeirotis. Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, 5(5):411–419, 2010.

[29] Eyal Peer, Joachim Vosgerau, and Alessandro Acquisti. Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, pages 1–9, 2013.

[30] Ivan P.L. Png, Chen-Yu Wang, and Qiu-Hong Wang. The deterrent and displacement effects of information security enforcement: International evidence. *Journal of Management Information Systems*, 25(2):125–144, 2008.

[31] Marc Riedel. Perceived circumstances, inferences of intent, and judgments of offense seriousness. *Journal of Law & Criminology*, 66(2):201–208, 1975.

[32] Julian V. Roberts. Public opinion, crime, and criminal justice. *Crime and Justice*, 16:99–180, 1992.

[33] Marcus K. Rogers, Kathryn Seigfried, and Kirti Tidke. Self-reported computer criminal behavior: A psychological analysis. *Digital Investigation*, 3, Supplement(0):116 – 120, 2006. The Proceedings of the 6th Annual Digital Forensic Research Workshop (DFRWS '06).

[34] Sean P. Rosenmerkel. Wrongfulness and harmfulness as components of seriousness of white-collar offenses. *Journal of Contemporary Criminal Justice*, 17(4):308–327, 2001.

[35] Joel Ross, Lilly Irani, M. Six Silberman, Andrew Zaldivar, and Bill Tomlinson. Who are the crowdworkers?: Shifting demographics in mechanical turk. In *CHI '10 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '10, pages 2863–2872, New York, NY, USA, 2010. ACM.

[36] Peter H. Rossi, Emily Waite, Christine E. Bose, and Richard E. Berk. The seriousness of crimes: Normative structure and individual differences. *American Sociological Review*, 39(2):224–237, 1974.

[37] Thorsten Sellin and M.E. Wolfgang. *The Measurement of Delinquency*. Wiley, 1964.

[38] Daniel J Simons and Christopher F Chabris. Common (mis)beliefs about memory: a replication and comparison of telephone and Mechanical Turk survey methods. *PloS One*, 7(12):e51876, 2012.

[39] H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly*, 20(2):167–196, 1996.

[40] Frank Stajano and Paul Wilson. Understanding scam victims: Seven principles for systems security. *Communications of the ACM*, 54(3):70–75, March 2011.

[41] Stelios Stylianou. Measuring crime seriousness perceptions: What have we learned and what else do we want to know. *Journal of Criminal Justice*, 1(31):37–56, 2003.

[42] Judith Jarvis Thomson. Killing, letting die, and the trolley problem. *The Monist*, 59(2):204–217, 1976.

[43] Monica A. Walker. Measuring the seriousness of crimes. *British Journal of Criminology*, 18(4):348–364, 1978.

[44] Mark Warr. What is the perceived seriousness of crimes? *Criminology*, 27(4):795–822, 1989.

A Appendix: Regression Tables

Table 2: Ordered probit regression results for the Type of Data experiment

	Wrongful	Harmful	Serious	Harsh	Sensitive	Respons.	Clever
Medical data	−0.112	0.188	0.073	−0.046	0.971***	0.025	0.003
	(0.140)	(0.145)	(0.147)	(0.143)	(0.151)	(0.154)	(0.142)
Female	0.467**	0.286	0.396**	0.125	0.380*	0.371*	−0.041
	(0.145)	(0.153)	(0.151)	(0.158)	(0.165)	(0.146)	(0.140)
US birth	−0.217	0.145	0.331	0.032	0.528	−0.356	0.167
	(0.231)	(0.346)	(0.302)	(0.298)	(0.327)	(0.458)	(0.211)
CFIP score	0.561***	0.194	0.294**	0.297**	0.491***	0.289*	0.241
	(0.107)	(0.115)	(0.101)	(0.101)	(0.116)	(0.117)	(0.131)
Freq. aff by cybercrime	−0.027	−0.011	−0.143	−0.094	−0.153	0.139	−0.294*
	(0.133)	(0.111)	(0.118)	(0.131)	(0.132)	(0.097)	(0.131)
Fake personal info	−0.010	−0.030	−0.046	−0.091	−0.085	0.040	−0.016
	(0.061)	(0.061)	(0.056)	(0.056)	(0.063)	(0.061)	(0.058)
Media awareness	−0.083	−0.026	0.010	0.015	0.034	0.077	0.062
	(0.051)	(0.048)	(0.048)	(0.043)	(0.051)	(0.051)	(0.053)
AC: Data	0.485	0.395	0.099	−0.103	0.054	−0.651*	−0.191
	(0.258)	(0.296)	(0.253)	(0.242)	(0.283)	(0.304)	(0.239)
AC: Context	−0.289	−0.295	−0.467**	−0.308*	−0.335*	0.095	0.151
	(0.166)	(0.178)	(0.164)	(0.148)	(0.163)	(0.173)	(0.165)
AC: Scope	−0.381	−0.569**	−0.501*	−0.176	−0.404	0.261	−0.127
	(0.195)	(0.209)	(0.212)	(0.185)	(0.230)	(0.199)	(0.208)
N	239	239	239	239	239	239	239
pseudo R^2	0.080	0.048	0.053	0.053	0.127	0.061	0.045

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the seven main Likert questions in the Type of Data experiment. The "Medical data" condition is versus the baseline condition of directory data. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 3: Ordered probit regression results for the Scope experiment

	Wrongful	Harmful	Serious	Harsh	Sensitive	Respons.	Clever
log(Num. Records)	0.069**	0.078**	0.159***	0.106***	0.135***	0.064*	0.058*
	(0.027)	(0.026)	(0.028)	(0.026)	(0.031)	(0.026)	(0.025)
Female	0.176	0.031	−0.015	0.097	0.241*	−0.142	0.101
	(0.097)	(0.096)	(0.095)	(0.092)	(0.111)	(0.094)	(0.094)
US birth	−0.239	0.045	−0.280	−0.308	−0.216	−0.007	−0.433
	(0.190)	(0.209)	(0.158)	(0.207)	(0.272)	(0.229)	(0.234)
CFIP score	0.357***	0.238***	0.378***	0.240***	0.630***	0.205**	0.261***
	(0.067)	(0.070)	(0.071)	(0.067)	(0.081)	(0.069)	(0.065)
Freq. aff by cybercrime	−0.096	−0.071	−0.187**	−0.101	−0.185*	−0.024	−0.025
	(0.064)	(0.060)	(0.063)	(0.062)	(0.076)	(0.062)	(0.063)
Fake personal info	0.042	−0.053	−0.023	−0.018	−0.030	0.012	0.064
	(0.041)	(0.039)	(0.039)	(0.038)	(0.045)	(0.040)	(0.040)
Media awareness	−0.040	−0.024	−0.030	−0.025	−0.037	0.049	−0.007
	(0.032)	(0.032)	(0.032)	(0.029)	(0.036)	(0.033)	(0.031)
AC: Data	−0.017	−0.127	−0.172	−0.115	0.037	0.011	0.331*
	(0.138)	(0.136)	(0.134)	(0.135)	(0.161)	(0.140)	(0.142)
AC: Context	0.023	−0.015	−0.089	−0.157	0.277	0.095	0.160
	(0.136)	(0.144)	(0.119)	(0.132)	(0.162)	(0.131)	(0.140)
AC: Scope	0.106	−0.032	0.076	0.259*	0.029	−0.047	0.218
	(0.127)	(0.133)	(0.128)	(0.131)	(0.151)	(0.127)	(0.140)
N	583	583	583	583	583	583	583
pseudo R^2	0.049	0.031	0.046	0.035	0.097	0.023	0.031

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the seven main Likert questions in the Scope experiment. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 4: Ordered probit regression results for the Motivation experiment (vs. Profiteer)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Student	-0.877^{***}	-0.323^{*}	-0.593^{***}	-0.791^{***}	-0.049	0.201	0.034	0.214
	(0.152)	(0.148)	(0.150)	(0.145)	(0.150)	(0.141)	(0.140)	(0.147)
Activist	-0.793^{***}	-0.252	-0.515^{***}	-0.485^{**}	-0.283	0.137	0.112	0.191
	(0.154)	(0.148)	(0.153)	(0.149)	(0.159)	(0.154)	(0.147)	(0.155)
Female	0.017	-0.025	0.063	-0.046	0.076	-0.112	0.352^{**}	-0.017
	(0.122)	(0.124)	(0.125)	(0.127)	(0.127)	(0.121)	(0.123)	(0.120)
US birth	-0.088	0.066	-0.058	0.330	0.040	-0.270	0.049	-0.336
	(0.213)	(0.259)	(0.226)	(0.253)	(0.277)	(0.237)	(0.320)	(0.251)
CFIP score	0.238^{**}	0.171	0.286^{**}	0.218^{*}	0.249^{**}	0.338^{***}	0.136	0.373^{***}
	(0.091)	(0.095)	(0.093)	(0.098)	(0.089)	(0.087)	(0.087)	(0.084)
Freq. aff by cybercrime	0.077	-0.054	0.107	-0.017	0.048	0.005	-0.129	-0.049
	(0.093)	(0.084)	(0.092)	(0.098)	(0.090)	(0.095)	(0.091)	(0.095)
Fake personal info	0.003	-0.006	0.054	-0.006	0.027	-0.028	0.060	-0.045
	(0.053)	(0.051)	(0.053)	(0.052)	(0.052)	(0.052)	(0.051)	(0.053)
Media awareness	0.012	0.101^{*}	0.054	0.032	0.101^{*}	0.028	0.033	-0.025
	(0.045)	(0.047)	(0.045)	(0.043)	(0.047)	(0.042)	(0.044)	(0.041)
AC: Data	-0.328^{**}	-0.123	-0.229	-0.289^{*}	-0.224	-0.521^{***}	0.064	-0.017
	(0.121)	(0.130)	(0.119)	(0.127)	(0.136)	(0.125)	(0.128)	(0.135)
AC: Context	0.047	0.209	0.031	0.095	-0.166	0.028	0.242	0.181
	(0.155)	(0.152)	(0.158)	(0.153)	(0.157)	(0.156)	(0.158)	(0.160)
AC: Scope	0.042	-0.091	0.112	-0.032	-0.070	0.119	-0.066	0.207
	(0.139)	(0.130)	(0.132)	(0.139)	(0.142)	(0.135)	(0.130)	(0.143)
AC: Motivation	-0.210	-0.128	-0.233	-0.229	-0.322	-0.565^{**}	-0.124	-0.017
	(0.180)	(0.192)	(0.177)	(0.170)	(0.188)	(0.179)	(0.174)	(0.191)
N	361	361	361	361	361	361	361	361
pseudo R^2	0.082	0.048	0.053	0.072	0.058	0.056	0.032	0.046

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Motivation experiment. The "Student" and "Activist" motivation conditions are versus the"Profiteer" baseline condition. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 5: Ordered probit regression results for the Motivation experiment (vs. Activist)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Student	−0.085	−0.071	−0.078	−0.306*	0.234	0.064	−0.078	0.024
	(0.150)	(0.139)	(0.149)	(0.150)	(0.153)	(0.146)	(0.148)	(0.145)
Profiteer	0.793***	0.252	0.515***	0.485**	0.283	−0.137	−0.112	−0.191
	(0.154)	(0.148)	(0.153)	(0.149)	(0.159)	(0.154)	(0.147)	(0.155)
Female	0.017	−0.025	0.063	−0.046	0.076	−0.112	0.352**	−0.017
	(0.122)	(0.124)	(0.125)	(0.127)	(0.127)	(0.121)	(0.123)	(0.120)
US birth	−0.088	0.066	−0.058	0.330	0.040	−0.270	0.049	−0.336
	(0.213)	(0.259)	(0.226)	(0.253)	(0.277)	(0.237)	(0.320)	(0.251)
CFIP score	0.238**	0.171	0.286**	0.218*	0.249**	0.338***	0.136	0.373***
	(0.091)	(0.095)	(0.093)	(0.098)	(0.089)	(0.087)	(0.087)	(0.084)
Freq. aff by cybercrime	0.077	−0.054	0.107	−0.017	0.048	0.005	−0.129	−0.049
	(0.093)	(0.084)	(0.092)	(0.098)	(0.090)	(0.095)	(0.091)	(0.095)
Fake personal info	0.003	−0.006	0.054	−0.006	0.027	−0.028	0.060	−0.045
	(0.053)	(0.051)	(0.053)	(0.052)	(0.052)	(0.052)	(0.051)	(0.053)
Media awareness	0.012	0.101*	0.054	0.032	0.101*	0.028	0.033	−0.025
	(0.045)	(0.047)	(0.045)	(0.043)	(0.047)	(0.042)	(0.044)	(0.041)
AC: Data	−0.328**	−0.123	−0.229	−0.289*	−0.224	−0.521***	0.064	−0.017
	(0.121)	(0.130)	(0.119)	(0.127)	(0.136)	(0.125)	(0.128)	(0.135)
AC: Context	0.047	0.209	0.031	0.095	−0.166	0.028	0.242	0.181
	(0.155)	(0.152)	(0.158)	(0.153)	(0.157)	(0.156)	(0.158)	(0.160)
AC: Scope	0.042	−0.091	0.112	−0.032	−0.070	0.119	−0.066	0.207
	(0.139)	(0.130)	(0.132)	(0.139)	(0.142)	(0.135)	(0.130)	(0.143)
AC: Motivation	−0.210	−0.128	−0.233	−0.229	−0.322	−0.565**	−0.124	−0.017
	(0.180)	(0.192)	(0.177)	(0.170)	(0.188)	(0.179)	(0.174)	(0.191)
N	361	361	361	361	361	361	361	361
pseudo R^2	0.082	0.048	0.053	0.072	0.058	0.056	0.032	0.046

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Motivation experiment. The "Student" and "Profiteer" motivation conditions are versus the"Activist" baseline condition. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 6: Ordered probit regression results for the Consequences experiment (vs. Low condition)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Acme High	0.183	0.408***	0.085	0.341**	0.152	0.001	−0.126	−0.025
	(0.123)	(0.122)	(0.119)	(0.123)	(0.136)	(0.140)	(0.116)	(0.118)
Customers High	0.056	0.377**	0.136	0.246*	0.101	0.056	0.105	−0.139
	(0.125)	(0.119)	(0.120)	(0.118)	(0.138)	(0.150)	(0.126)	(0.124)
Female	0.172	0.114	0.170	0.167	0.258*	0.205	0.115	0.069
	(0.106)	(0.103)	(0.100)	(0.101)	(0.116)	(0.121)	(0.106)	(0.103)
US birth	0.045	−0.117	0.151	0.099	−0.002	−0.145	0.084	−0.078
	(0.239)	(0.216)	(0.241)	(0.237)	(0.222)	(0.269)	(0.286)	(0.215)
CFIP score	0.206**	0.168*	0.291***	0.160*	0.422***	0.656***	0.227**	0.125
	(0.076)	(0.082)	(0.078)	(0.080)	(0.100)	(0.103)	(0.074)	(0.078)
Freq. aff by cybercrime	−0.016	−0.002	−0.032	0.011	−0.018	−0.101	0.007	0.016
	(0.079)	(0.076)	(0.078)	(0.077)	(0.088)	(0.097)	(0.075)	(0.073)
Fake personal info	−0.108*	−0.054	−0.093*	−0.114**	−0.029	−0.002	0.089*	0.016
	(0.043)	(0.043)	(0.043)	(0.042)	(0.052)	(0.047)	(0.043)	(0.041)
Media awareness	−0.042	0.028	−0.027	−0.019	0.078	0.058	0.039	0.024
	(0.039)	(0.037)	(0.038)	(0.039)	(0.045)	(0.048)	(0.041)	(0.037)
AC: Data	0.400*	0.137	0.240	0.203	0.315	0.446*	0.331*	0.096
	(0.157)	(0.140)	(0.145)	(0.146)	(0.169)	(0.187)	(0.142)	(0.147)
AC: Context	−0.101	−0.068	0.035	−0.071	−0.054	−0.019	−0.325**	−0.114
	(0.129)	(0.114)	(0.126)	(0.120)	(0.139)	(0.149)	(0.124)	(0.130)
AC: Scope	0.005	−0.110	−0.097	−0.108	−0.077	0.114	0.095	0.127
	(0.108)	(0.110)	(0.111)	(0.109)	(0.125)	(0.133)	(0.113)	(0.109)
AC: Consequence	−0.071	−0.119	−0.127	−0.196	−0.164	−0.154	0.193	0.060
	(0.170)	(0.203)	(0.185)	(0.188)	(0.215)	(0.204)	(0.166)	(0.178)
N	479	479	479	479	479	479	479	479
pseudo R^2	0.045	0.034	0.040	0.039	0.079	0.119	0.033	0.015

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Consequences experiment. The "Acme High" and "Customers High" motivation conditions are the conditions in which Acme was described as experiencing high losses and its customers were described as experiencing high losses, respectively. Both rare versus the"Low" baseline condition in which Acme was described as experiencing minimal losses. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 7: Ordered probit regressions for the Consequences experiment (vs. Customers)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Low	−0.056	−0.377**	−0.136	−0.246*	−0.101	−0.056	−0.105	0.139
	(0.125)	(0.119)	(0.120)	(0.118)	(0.138)	(0.150)	(0.126)	(0.124)
Acme High	0.126	0.030	−0.051	0.094	0.051	−0.055	−0.231	0.114
	(0.130)	(0.122)	(0.124)	(0.122)	(0.139)	(0.154)	(0.119)	(0.125)
Female	0.172	0.114	0.170	0.167	0.258*	0.205	0.115	0.069
	(0.106)	(0.103)	(0.100)	(0.101)	(0.116)	(0.121)	(0.106)	(0.103)
US birth	0.045	−0.117	0.151	0.099	−0.002	−0.145	0.084	−0.078
	(0.239)	(0.216)	(0.241)	(0.237)	(0.222)	(0.269)	(0.286)	(0.215)
CFIP score	0.206**	0.168*	0.291***	0.160*	0.422***	0.656***	0.227**	0.125
	(0.076)	(0.082)	(0.078)	(0.080)	(0.100)	(0.103)	(0.074)	(0.078)
Freq. aff by cybercrime	−0.016	−0.002	−0.032	0.011	−0.018	−0.101	0.007	0.016
	(0.079)	(0.076)	(0.078)	(0.077)	(0.088)	(0.097)	(0.075)	(0.073)
Fake personal info	−0.108*	−0.054	−0.093*	−0.114**	−0.029	−0.002	0.089*	0.016
	(0.043)	(0.043)	(0.043)	(0.042)	(0.052)	(0.047)	(0.043)	(0.041)
Media awareness	−0.042	0.028	−0.027	−0.019	0.078	0.058	0.039	0.024
	(0.039)	(0.037)	(0.038)	(0.039)	(0.045)	(0.048)	(0.041)	(0.037)
AC: Data	0.400*	0.137	0.240	0.203	0.315	0.446*	0.331*	0.096
	(0.157)	(0.140)	(0.145)	(0.146)	(0.169)	(0.187)	(0.142)	(0.147)
AC: Context	−0.101	−0.068	0.035	−0.071	−0.054	−0.019	−0.325**	−0.114
	(0.129)	(0.114)	(0.126)	(0.120)	(0.139)	(0.149)	(0.124)	(0.130)
AC: Scope	0.005	−0.110	−0.097	−0.108	−0.077	0.114	0.095	0.127
	(0.108)	(0.110)	(0.111)	(0.109)	(0.125)	(0.133)	(0.113)	(0.109)
AC: Consequence	−0.071	−0.119	−0.127	−0.196	−0.164	−0.154	0.193	0.060
	(0.170)	(0.203)	(0.185)	(0.188)	(0.215)	(0.204)	(0.166)	(0.178)
N	479	479	479	479	479	479	479	479
pseudo R^2	0.045	0.034	0.040	0.039	0.079	0.119	0.033	0.015

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Consequences experiment. The "Acme High" and "Low" motivation conditions are versus the"Customer High" baseline condition. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 8: Ordered probit regressions for the Co-Responsibility experiment

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Not Patched	0.132	0.099	0.155	0.074	0.089	−0.364*	0.420**	−0.184
	(0.136)	(0.136)	(0.133)	(0.132)	(0.150)	(0.163)	(0.129)	(0.136)
Female	0.190	0.173	0.136	0.058	0.162	−0.032	0.131	0.029
	(0.148)	(0.154)	(0.152)	(0.144)	(0.160)	(0.175)	(0.142)	(0.150)
US birth	0.225	−0.757**	0.368	0.228	−0.345	−0.599	−0.484	0.213
	(0.356)	(0.288)	(0.272)	(0.235)	(0.428)	(0.437)	(0.366)	(0.394)
CFIP score	0.576***	0.395**	0.558***	0.385***	0.700***	1.080***	0.251*	0.364**
	(0.119)	(0.129)	(0.116)	(0.112)	(0.137)	(0.142)	(0.113)	(0.126)
Freq. aff by cybercrime	−0.004	0.033	−0.029	−0.049	0.033	0.004	0.097	0.005
	(0.083)	(0.089)	(0.103)	(0.092)	(0.107)	(0.117)	(0.099)	(0.098)
Fake personal info	−0.015	−0.152*	0.001	−0.121	0.024	−0.010	−0.049	0.066
	(0.060)	(0.064)	(0.065)	(0.063)	(0.070)	(0.065)	(0.061)	(0.070)
Media awareness	0.030	0.112*	0.093	0.076	0.042	−0.066	0.176**	0.064
	(0.048)	(0.050)	(0.049)	(0.048)	(0.056)	(0.062)	(0.054)	(0.054)
AC: Data	−0.267	−0.286	−0.335	−0.057	−0.076	−0.214	−0.180	0.182
	(0.205)	(0.215)	(0.187)	(0.219)	(0.208)	(0.272)	(0.197)	(0.200)
AC: Context	−0.356*	−0.351*	−0.285	−0.233	−0.552**	−0.249	−0.135	0.076
	(0.161)	(0.149)	(0.161)	(0.148)	(0.179)	(0.191)	(0.163)	(0.170)
AC: Scope	0.004	0.217	0.262	0.080	0.498**	0.387*	0.206	0.203
	(0.190)	(0.177)	(0.171)	(0.161)	(0.193)	(0.174)	(0.179)	(0.169)
AC: Patched	−0.287	−0.325	−0.185	−0.282	−0.279	−0.471	−0.349	0.034
	(0.233)	(0.227)	(0.210)	(0.210)	(0.287)	(0.282)	(0.224)	(0.240)
N	276	276	276	276	276	276	276	276
pseudo R^2	0.060	0.054	0.052	0.039	0.107	0.165	0.058	0.050

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Co-responsibility experiment. The "Not Patched" condition is versus the "Patched" baseline condition in which Acme was described as having patched its servers. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 9: Ordered probit regressions for the Context experiment (vs. Bank)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Government	−0.063	0.005	−0.035	−0.036	0.141	−0.126	0.155	−0.031
	(0.119)	(0.120)	(0.125)	(0.116)	(0.139)	(0.141)	(0.118)	(0.116)
Non-Profit	0.045	−0.030	−0.224	0.028	0.096	−0.211	−0.359**	−0.189
	(0.122)	(0.124)	(0.122)	(0.121)	(0.141)	(0.155)	(0.120)	(0.121)
Org. size	0.055	0.064	0.045	0.053	0.133**	0.148**	0.059	0.142**
	(0.044)	(0.041)	(0.043)	(0.043)	(0.047)	(0.050)	(0.046)	(0.046)
Female	0.005	0.003	−0.041	−0.016	0.090	0.069	0.155	0.131
	(0.102)	(0.101)	(0.100)	(0.099)	(0.116)	(0.118)	(0.096)	(0.100)
US birth	−0.080	−0.102	0.061	−0.166	−0.301	0.156	0.123	−0.065
	(0.277)	(0.273)	(0.223)	(0.248)	(0.284)	(0.376)	(0.299)	(0.276)
CFIP score	0.356***	0.194*	0.378***	0.209**	0.406***	0.520***	0.139	0.136
	(0.073)	(0.076)	(0.073)	(0.079)	(0.085)	(0.076)	(0.075)	(0.073)
Freq. aff by cybercrime	−0.023	−0.033	−0.057	−0.046	−0.116	0.002	−0.029	0.051
	(0.065)	(0.063)	(0.064)	(0.060)	(0.073)	(0.077)	(0.064)	(0.069)
Fake personal info	−0.023	0.002	−0.018	−0.008	0.051	−0.079	0.014	−0.011
	(0.041)	(0.040)	(0.040)	(0.037)	(0.046)	(0.044)	(0.042)	(0.040)
Media awareness	−0.026	−0.047	0.030	−0.010	−0.030	−0.011	0.065	0.063
	(0.036)	(0.038)	(0.036)	(0.035)	(0.044)	(0.043)	(0.038)	(0.038)
AC: Data	0.029	0.029	0.038	0.021	0.374*	0.376*	−0.181	−0.054
	(0.152)	(0.151)	(0.153)	(0.152)	(0.160)	(0.156)	(0.124)	(0.142)
AC: Context	0.007	0.078	−0.185	0.018	−0.018	−0.029	−0.103	−0.145
	(0.129)	(0.131)	(0.127)	(0.133)	(0.145)	(0.161)	(0.123)	(0.128)
AC: Scope	−0.157	0.044	−0.107	−0.039	−0.028	0.165	−0.024	0.036
	(0.123)	(0.137)	(0.127)	(0.124)	(0.144)	(0.143)	(0.127)	(0.127)
N	502	502	502	502	502	502	502	502
pseudo R^2	0.043	0.022	0.045	0.029	0.073	0.092	0.028	0.033

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Context experiment. The "Government" and "Non-profit" conditions are versus the "Bank" baseline condition. Regressions also included categorical control variables for occupation, age, education, and work situation.

Table 10: Ordered probit regressions for the Context experiment (vs. Non-Profit)

	Wrongful	Harmful	Serious	Harsh	Pot. Harm.	Sensitive	Respons.	Clever
Bank	−0.045	0.030	0.224	−0.028	−0.096	0.211	0.359**	0.189
	(0.122)	(0.124)	(0.122)	(0.121)	(0.141)	(0.155)	(0.120)	(0.121)
Government	−0.108	0.035	0.189	−0.064	0.045	0.085	0.513***	0.158
	(0.116)	(0.118)	(0.112)	(0.112)	(0.137)	(0.132)	(0.118)	(0.119)
Org. size	0.055	0.064	0.045	0.053	0.133**	0.148**	0.059	0.142**
	(0.044)	(0.041)	(0.043)	(0.043)	(0.047)	(0.050)	(0.046)	(0.046)
Female	0.005	0.003	−0.041	−0.016	0.090	0.069	0.155	0.131
	(0.102)	(0.101)	(0.100)	(0.099)	(0.116)	(0.118)	(0.096)	(0.100)
US birth	−0.080	−0.102	0.061	−0.166	−0.301	0.156	0.123	−0.065
	(0.277)	(0.273)	(0.223)	(0.248)	(0.284)	(0.376)	(0.299)	(0.276)
CFIP score	0.356***	0.194*	0.378***	0.209**	0.406***	0.520***	0.139	0.136
	(0.073)	(0.076)	(0.073)	(0.079)	(0.085)	(0.076)	(0.075)	(0.073)
Freq. aff by cybercrime	−0.023	−0.033	−0.057	−0.046	−0.116	0.002	−0.029	0.051
	(0.065)	(0.063)	(0.064)	(0.060)	(0.073)	(0.077)	(0.064)	(0.069)
Fake personal info	−0.023	0.002	−0.018	−0.008	0.051	−0.079	0.014	−0.011
	(0.041)	(0.040)	(0.040)	(0.037)	(0.046)	(0.044)	(0.042)	(0.040)
Media awareness	−0.026	−0.047	0.030	−0.010	−0.030	−0.011	0.065	0.063
	(0.036)	(0.038)	(0.036)	(0.035)	(0.044)	(0.043)	(0.038)	(0.038)
AC: Data	0.029	0.029	0.038	0.021	0.374*	0.376*	−0.181	−0.054
	(0.152)	(0.151)	(0.153)	(0.152)	(0.160)	(0.156)	(0.124)	(0.142)
AC: Context	0.007	0.078	−0.185	0.018	−0.018	−0.029	−0.103	−0.145
	(0.129)	(0.131)	(0.127)	(0.133)	(0.145)	(0.161)	(0.123)	(0.128)
AC: Scope	−0.157	0.044	−0.107	−0.039	−0.028	0.165	−0.024	0.036
	(0.123)	(0.137)	(0.127)	(0.124)	(0.144)	(0.143)	(0.127)	(0.127)
N	502	502	502	502	502	502	502	502
pseudo R^2	0.043	0.022	0.045	0.029	0.073	0.092	0.028	0.033

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: The table shows ordered probit regression results for responses to the eight main Likert questions in the Context experiment. The "Bank" and "Government" conditions are versus the "Non-profit" baseline condition. Regressions also included categorical control variables for occupation, age, education, and work situation.

B Appendix: Experiment Design

B.1 Scenario Presentation

We presented participants with a page of instructions, followed by a page presenting a scenario describing a data breach cybercrime committed by Tom Smith. In all experiments except the Context experiment, the scenario described Tom Smith as breaking into the Acme Insurance Company (in the Context experiment, the company was called "ACR").

B.2 Memory Checks

After they read the scenario, they survey presented participants with three or four questions about the scenario intended to test their recall of (and attention to) the details of that scenario. Each experiment included the following three questions. Possible answers varied between experiments depending on the scenario. Dummy variable names indicating whether the question was answered correctly are listed after each question:

- "Which one of the following kinds of data was among the data Tom Smith accessed?" [Health histories, Addresses Credit card numbers, Social security numbers, Driver's license numbers, Passwords, Account numbers, None of the above] (`ac_data_ok`)
- "What did Tom Smith break into?" [Hospital, Retailer, Bank, Manufacturer, School, Insurer, None of the above] (`ac_context_ok`)
- "How many records did Tom Smith download?" [1,000, 10,000, 100,000, 1,000,000, None of the above] (`ac_scope_ok`)

If the manipulation was not covered by one of these three questions, we added an additional question to check recall of the manipulation. After each memory-check question, the survey presented each participant with a page indicating whether his or her answer was correct. In either case, the correct answer was repeated. This was an attempt to further reinforce the participant's awareness of the details of the scenario.

B.3 Main Likert Questions

Following the memory check questions, they survey asked participants to answer several questions on a 1–7 Likert scale. The questions were presented in random order. Variable names are listed after each question:

- "How wrongful were Tom Smith's actions?" (`how_wrongful`)
- "How serious was the crime Tom Smith committed?" (`how_serious`)
- "How harshly should Tom Smith be punished?" (`how_harshly`)
- "How harmful were Tom Smith's actions?' (`how_harmful`)
- "How responsible was the Acme Insurance Company for the crime?" (`how_responsible`)[6]
- "How clever was Mr. Tom Smith?" (`how_clever`)
- "How sensitive was the data that Tom Smith downloaded?" (`how_sensitive`)

[6]In the Context experiment, the"Acme Insurance Company" was replaced by "ACR."

B.4 Specific Punishment

Participants were then asked to recommend a specific punishment for Tom Smith. This was done to obtain a measure of how harshly a crime should be punished that was closer to an interval scale than ordinal. The question was the following:

- "Please assume that Tom Smith is convicted in a court of law. How long of a jail or prison sentence do you believe would most appropriately fit the crime? Please assume that the entire term of a sentence would be served." (`sentence`)
  - No probation, jail, or prison time
  - Probation only
  - 0 to 29 days served in jail or prison
  - 30 to 89 days served in jail or prison
  - 90 to 179 days served in jail or prison
  - 180 to 364 days served in jail or prison
  - 1 year to less than 2 years served in jail or prison
  - 2 years to less than 5 years served in jail or prison
  - 5 years to less than 10 years served in jail or prison
  - 10 years to less than 20 years served in jail or prison
  - 20 or more years served in jail or prison
  - Life served in jail or prison

B.5 Potential Consequences

The Motivation, Consequences, Co-Responsibility, and Context experiments followed the specific-punishment question with a question about the *potential* consequences of Tom Smith's actions. This question was intended to help determine whether participants might be judging scenarios by potential consequences instead of the actual consequences that were described in the scenarios. The added question also made an additional attention check possible: participants who rated the potential consequences as lower than the actual consequences may not have been paying enough attention to the questions.

- "How harmful might the potential consequences of Tom Smith's actions have been?" (`how_pot_harmful`)

This question was placed after the main Likert questions and the specific-punishment question to avoid priming participants to think in terms of potential instead of actual consequences.

B.6 Privacy Attitudes

The next section of the survey included several questions intended to measure participants' attitudes about privacy. This used the 15-question Concern for Information Privacy (CFIP) scale. Each question in the scale is answered on 1–7 Likert scale from "Strongly disagree" to "Strongly agree." These questions were randomized.

- "Here are some statements about personal information. From the standpoint of personal privacy, please indicate the extent to which you, as an individual, agree or disagree with each statement."
  A. "It usually bothers me when companies ask me for personal information."

B. "All the personal information in computer databases should be double-checked for accuracy—no matter how much this costs."
C. "Companies should not use personal information for any purpose unless it has been authorized by the individuals who provided the information."
D. "Companies should devote more time and effort to preventing unauthorized access to personal information."
E. "When companies as me for personal information, I sometimes think twice before providing it."
F. "Companies should take more steps to make sure that the personal information in their files is accurate."
G. "When people give personal information to a company for some reason, the company should never use the information for any other reason."
H. "Companies should have better procedures to correct errors in personal information."
I. "Computer databases that contain personal information should be protected from unauthorized access—no matter how much it costs."
J. "It bothers me to give personal information to so many companies."
K. "Companies should never sell the personal information in their computer databases to other companies."
L. "Companies should devote more time and effort to verifying the accuracy of the personal information in their databases."
M. "Companies should never share personal information with other companies unless it has been authorized by the individuals who provided the information."
N. "Companies should take more steps to make sure that unauthorized people cannot access personal information in their computers."
O. "I'm concerned that companies are collecting too much personalization about me."

Items A, E, J, and O comprise the "Collection" subscale; items B, F, H, and L comprise the "Errors" subscale; items C, G, K, and M comprise the "Unauthorized Secondary Use" subscale; and items D, I, and N comprise the "Improper Access" subscale. An overall score (`CFIP_overall`) is calculated by averaging the sub-scale scores.

Three other privacy attitude questions were asked:

- "How frequently have you personally been the victim of cybercrime or an invasion of privacy?" (`aff_by_cybercrime`) [Never, Once, A few times, Several times]
  - "If you have been a victim of a cybercrime or invasion of privacy, can you please provide more details?" (`aff_by_cybercrime_text`)
- "Some websites ask you to register with the site by providing personal information. When asked for such information, how often do you provide incorrect information?" (`fake_personal_info`) [I have never given incorrect information, Under 25% of the time, 26%–50% of the time, 51%–75% of the time, 76% or more of the time]
- "How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet?" (`media_awareness`) [Likert; 1=Not at all, 7=Very much]

B.6.1 Demographics

Participants were asked the following demographic questions:

- "What is your gender?" (`gender`)

- "What is your age?" [<18, 18–24, 25–34, 35–44, 45–54, 55–64, 65+] (`age_category`)
- "What country were you born in?" (`residence_co`)
- "What country do you live in now?" (`birth_co`)
- "What is the highest level of education you have completed?" [Less than High School, High School/GED, Some college, 2-year college degree, 4-year college degree, Masters degree, Doctoral degree, Professional degree (JD/MD)] (`education`)
- "What is your current work situation?" [Student, Stay at home caregiver or homemaker, Unemployed, Part-time employed, Full-time employed, Self-employed, Retired, Not working for some other reason (e.g., disability, illness, etc.)] (`work_situation`)
- "Which of the following most closely describes your occupation?"(`occupation`)

Each demographic question except for the two country questions included the option "Prefer not to answer."

B.7 Open-Ended Questions

Finally, the survey asked four open-ended questions:

- "What do you think of Tom Smith?" (`tom_opinion_text`)
- "What do you think the *actual* consequences of Tom Smith's actions were?"[7] (`actual_conseq_text`)
- "What do you think the potential consequences of Tom Smith's actions could have been?"[8] (`potential_conseq_text`)
- "Finally, what do you think this study was about?" (`study_about_text`)

B.8 Data Cleaning

The data cleaning process included the following steps:

1. Deleting unused or unneeded columns.
2. Removing incomplete entries. Entries that had a "finished" value of 0 were not removed if all answers were completed.
3. Removing duplicates, as determined by duplicate MTurk ID or IP address
4. Removing responses in which the participant indicated that he or she resided outside the United States. The MTurk assignment was set only to be available to workers who lived in the United States.
5. Removing any responses from participants who rated the scenario as more harmful than potentially harmful, a pairing of answers that suggests lack of attention.
6. Checking open responses for signs of unreliable responses.

Note that responses from participants who answered one or more memory check question incorrectly were *not* removed. Accuracy in answering the memory check questions is controlled for in the regressions.

[7] This question was included in the Motivation, Consequences, Co-Responsibility, and Context experiments.
[8] This question was included in the Motivation, Consequences, Co-Responsibility, and Context experiments.

B.9 Dummy Variables

We created the following dummy variables for data analysis in all experiments:

- `female`: Set to 1 if `gender` = 1, 0 otherwise.
- `us_residence`: Set to 1 if the `residence_co` is the U.S., 0 otherwise.
- `us_birth`: Set to 1 if `birth_co` is the U.S., 0 otherwise.