Evaluating Behavioral Interventions for Privacy Decisions

Submitted in partial fulfillment of the requirements for

the degree of

Doctor of Philosophy

in

Engineering and Public Policy

Logan Warberg

B.S. Computer Science, Montana State University
B.A. Political Science, Montana State University
M.S. Engineering and Public Policy, Carnegie Mellon University

Carnegie Mellon University
Pittsburgh, PA

August 2022

Acknowledgements

Over the course of completing this work, I have benefited from the guidance and feedback of my committee members: Dr. Alessandro Acquisti, Dr. Douglas Sicker, Dr. Alex Davis, and Dr. Cristobal Cheyre. Each played a pivotal role in my development as a researcher over the past few years. I'd like to thank Dr. Cristobal Cheyre for helping me think through the many modeling problems in my final chapter. I'd also like to thank Dr. Alex Davis for helping me understand regression methods and instilling in me the importance of pre-registrations and study reproducibility. Finally, I'd like to thank my advisors and co-chairs Dr. Alessandro Acquisti and Dr. Douglas Sicker for guiding me through the PhD and showing me how design experiments and conduct sound research. Their advice over countless hours of meetings and emails were essential to me making it to this point.

In addition to my committee, I'd like to thank mentors and collaborators including Dr. Baruch Fischhoff, Dr. Norman Sadeh, and Dr. Lorrie Cranor for their insight and advice.

Beyond the countless hours of research, my time in Pittsburgh was made better by the company of many friends. Weekly meetings of the 'Breakfast Buddies', the EPP Dungeons and Dragons group, and 'Film Club' provided needed distraction from research, and later, needed social interaction during COVID isolation. I'd like to thank Tobi Adekanye, Kristen Allen, Matt Babcock, Vanya Britto, Jessica Colnago, Christophe Combemale, Daniel Gringerich, Mike Rath, Danyelle & Pierce Ware, and many more for their support and continued friendship.

Most importantly, I would like to express my gratitude to my parents Lori and Will Warberg for their continuous support over the past seven years and being with me though each high and low that came with completing the PhD. Without them, this work would not have been possible. Finally, I'd like to thank my grandparents, aunts, uncles, and cousins who provided encouragement and motivation to help me complete this dissertation.

This research was made possible in part through funding from the Department of Engineering and Public Policy, DARPA and the Air Force Research Laboratory, as part of the DARPA Brandeis program, under grant number FA8750-15-2-0277, the United States Department of Education under grant number P200A190017, Google LLC, and Samsung Research America under grant number IC-5010.

Abstract

Behavioral interventions have been developed to help users overcome structural and cognitive hurdles in privacy decisions by modifying them to add new information or change how a decision is presented. In the hands of practitioners and policy-makers, behavioral interventions are a tool which can be used to potentially reduce the complexity of privacy decisions for users, leading to better outcomes. This dissertation explores the application of different behavioral interventions to contexts currently unexplored in the literature and examines whether they can be made more effective. In addition, this dissertation contributes to policy discussions regarding behavioral interventions for privacy decisions by examining how data protection regulation has impacted the presence and strength of behavioral interventions on websites.

In Chapter 2 of this dissertation, I present the results of three online experiments aimed at investigating whether nudges (a type of behavioral intervention) tailored to various psychometric scales can influence participants' disclosure choices. Across each of these experiments, we do not find significant effects robustly linking any of the measured psychometric variables to differences in disclosure rates. In Chapter 3, I examine whether providing explanations (a different type of behavioral intervention) for recommendations about hypothetical privacy decisions impacts the rate at which participants adopt the recommendations. The results do not find strong evidence linking the tested explanations to hypothetical adoption intent. Across both chapters, I describe the study designs and contextualize the null results, drawing lessons about the robustness of behavioral interventions in the context of real user behaviors.

In Chapter 4, I leverage longitudinal data from a large panel of websites to investigate the dynamics of privacy notice design over time in order to comment on the potential impacts of the EU's General Data Protection Regulation on the presence and strength of behavioral interventions within privacy notices. The results identify geographic disparities in the prevalence of privacy notices in the US compared to the EU. I additionally identify a shift away from more coercive responses on US websites and a shift towards consent responses on EU websites. I discuss the implications of these results, limitations, and possible directions for future research.

Finally, in Chapter 5, I conclude this dissertation by discussing considerations practitioners should take into account when considering behavioral interventions for privacy decisions and lessons for policy-makers regarding the use of behavioral interventions in public policy.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

As technology has grown in complexity, so have the decisions that users must navigate to manage their privacy. These decisions can often be difficult to make due to structural and cognitive hurdles that hinder the ability of users to properly assess the risks of information disclosure [1]. Structural hurdles such as the inability for users to conceive of all possible future risks at the time of disclosure prevent users from making informed choices, while cognitive hurdles such as heuristics and biases can lead to decisions that result in privacy harms [5]. Both types of hurdles come into play when users are faced with privacy policies which can be lengthy and written in a manner which is difficult for users to understand [13]. With the increasing capability of firms and organizations to draw inferences between disparate pieces of information, privacy harms may arise [93].

Behavioral interventions can help users overcome structural and cognitive hurdles in privacy decisions by modifying them to add new information or change how a decision is presented. Together, behavioral interventions offer one tool for practitioners and policy-makers to modify privacy decisions in order to reduce their complexity for users. Behavioral interventions may attempt to resolve information asymmetries or correct for biases by re-framing decisions or adding additional context. In this space, nudges have been shown to be a powerful tool [98]. Applied to privacy, nudges can impact a decision-making process wherein privacy preferences are often contextual and can be incomplete or poorly defined [3].

Oftentimes, nudges can take a paternalistic approach to behavioral interventions - where the nudge leverages biases to guide users toward particular outcomes. For example, the pre-selection of a default option in a decision can potentially be seen as a nudge which applies status-quo bias to encourage users to choose the pre-selected option [4]. A different approach to help users navigate privacy decisions is through interventions which provide additional context and information to decisions, but do not prescribe an outcome. Both types of interventions are of potential interest to practitioners and policy-makers.

To contribute to the ongoing discourse on the use of behavioral interventions in privacy disclosure decisions, this dissertation explores different types of interventions by applying them to contexts currently unexplored in the literature and examining whether they can be made more effective. In addition, this dissertation contributes to policy discussions regarding behavioral interventions by examining how data protection regulation has impacted the presence and design of behavioral interventions on websites.

Specifically, this thesis addresses the following broad research questions:

1. Can privacy nudges for disclosure decisions be tailored to individual differences in personality traits and decision making ability?

2. Do different explanations for hypothetical mobile privacy recommendations impact users' adoption intentions?

3. How has the design and presence of behavioral interventions within privacy notices on websites changed over time following the implementation of the European Union's General Data Protection Regulation (GDPR)?

The following chapters of this thesis examine each of the above research questions through a combination of controlled experiments and an observational study. In Chapter 2, I conduct three experimental studies to determine whether privacy nudges can be tailored to individual differences in personality traits and decision making ability. Chapter 3 details an experiment designed to test the effectiveness of different styles of explanations for mobile privacy recommendations. Together, the studies in these two chapters examine two different types of behavioral interventions in different contexts. In Chapter 4, I utilize observational data from a longitudinal panel of websites to model changes in privacy notice design following the enforcement of the GDPR. Finally, in Chapter 5 I examine the results of the preceding chapters together in order to draw lessons for choice architects and policy makers regarding the design and application of behavioral interventions for privacy decisions.

Chapter 2

Tailoring Privacy Nudges to Individual Differences

Abstract

While the effectiveness of nudges in influencing user behavior has been documented within the literature, most prior work in the privacy field has focused on ‘one-size-fits-all’ interventions. Recent behavioral research has identified the potential of tailoring nudges to users by leveraging individual differences in decision making and personality. We present the results of three online experiments aimed at investigating whether nudges tailored to various psychometric scales can influence participants’ disclosure choices. Each study adopted a difference-in-differences design, testing whether differences in disclosure rates for participants presented with a nudge were affected by differences along various psychometric variables. Study 1 used a hypothetical disclosure scenario to measure participants’ responses to a single nudge. Study 2 and its replication (Study 3) tested responses in real disclosure scenarios to two nudges. Across all studies, we failed to find significant effects robustly linking any of the measured psychometric variables to differences in disclosure rates. We describe our study design and results along with a discussion of the practicality of using decision making and personality traits to tailor privacy nudges.

2.1 Introduction

Nudges have emerged within the privacy and security literature as an effective means of affecting, and possibly assisting, user behavior [4]. Nudges work by modifying the structure of choices to encourage certain behaviors without altering economic incentives [98]. Recent research has applied nudges to diverse privacy and security scenarios where users face hurdles in the decision making process, and where those hurdles can result in negative outcomes. Within the security literature, nudges have been used to steer users towards better security behaviors in areas including password creation [103], heeding browser warnings [66], and selecting wireless networks [102]. In a similar vein, nudges within the privacy literature have been used to guide users to better privacy choices, often by providing them with salient information to aid the decision making process regarding online disclosures. Privacy nudges have been used to help add context to decisions such as setting mobile app permissions [9] and posting information on social networks [107]. Other nudges have focused on changes in the presentation of choices to influence user behavior; framing effects have been shown to encourage users to allow or prohibit information disclosures [85].

While effective, many of the applications of nudges within the privacy literature have focused on 'one-size-fits-all' approaches, where a certain behavioral intervention is applied to a diverse set of individuals, with no tailoring of the nudge based on characteristics unique to each individual. This is starting to change, where recent studies have examined nudges applied to privacy behaviors that are tailored to traits such as demographics [52]. Recent behavioral work has highlighted the possibility of making nudges more effective by tailoring them based on individual traits [30]. Given the prominent role that differences in individual traits can play in terms of privacy decision making, one could expect the privacy field to be one where the tailoring of nudges could prove particularly effective. Previous work has identified significant diversity in the privacy attitudes and behaviors across cultures and among different individuals within the same culture [72, 105, 109]. Existing research has also examined individual psychometric differences and found them to be predictors of differences in privacy attitudes [30]. This diversity suggests that people may approach privacy decisions differently. Consequently, behavioral interventions that capture such diversity may be more effective in changing user behavior. Specifically, differences in decision making and personality may have applications in creating personalized or 'tailored' nudges. In the security field, recent work has started investigating the effectiveness of 'tailored' nudges in influencing security behavior [66, 82].

We investigate psychometrically tailored nudges in the context of privacy behavior. We conducted three online experiments that attempted to identify effects for tailored nudges on data disclosure choices. Across the three studies, participants completed surveys in which they were presented with a hypothetical (Study 1) or real (Studies 2 and 3) disclosure choice. Each disclosure choice was paired with a nudge designed to encourage participants to either allow or prohibit the disclosure of potentially sensitive information. After recording participants' disclosure choices, we measured participants along a variety of psychometric scales designed to capture individual differences in decision making and personality. We examined whether changes in the differences in disclosure rates could be predicted by the measured psychometric variables. We modeled this relationship using logistic regression where the disclosure choice made by participants was our dependent variable.

In Study 1, we presented participants with a hypothetical disclosure choice and a single framing nudge with 'Opt-In' and 'Opt-Out' conditions. We examined the interaction between these conditions and two psychometric variables. Our analysis found a significant main effect for the nudge, but failed to identify significant effects for the *interaction* between the nudge and participants' psychometric traits. In other words: the nudge was effective in influencing participants' disclosure choice, but its effectiveness did not vary with participants' psychometric traits.

Lessons from Study 1 informed the design of Study 2. This study explored a wider scope of nudges and psychometric variables (15 psychometric variables across 2 nudges, across 4 experimental conditions) in a real disclosure setting. In our analysis of Study 2, we found - again - a main effect for the nudges. We additionally identified three potentially significant effects for the interaction between nudge conditions and psychometric variables. Thus, Study 2 suggested that the effectiveness of some nudges could in fact vary with participants' psychometric traits.

Building on the results of Study 2, we conducted a replication study (Study 3) to test the robustness of the effects found in Study 2. Study 3 used the same design and experimental conditions as Study 2. However, our analysis of Study 3 failed to replicate the effects we identified in Study 2.

We interpret the null findings (and in particular the failure to replicate the initial significant findings) as a cautionary tale for both the practical effectiveness of tailored nudges and for future research conducted in this area. The results suggest first, that while one cannot exclude on statistical grounds the theoretical possibility of tailoring nudges to individual differences in decision making and personality, the effects of some of these interventions may be fragile, and potentially impractical for many applications. Second, given the risk of spurious correlations emerging as significant from the interaction of multiple nudges with multiple psychometric variables, research in this area should pay particular attention to replicating and validating results of such interactions.

2.2 Related Work

Our work builds upon the existing body of research on nudging [98]. As defined by Thaler and Sunstein, nudges are changes in the design or structure of a choice which predictably alter behavior without altering economic incentives. Building upon this idea, studies within the psychology and behavioral economics literature have explored and expanded upon nudge interventions, becoming a popular form of behavioral intervention [38].

2.2.1 Nudges in Privacy and Security

Within privacy and security, nudges have grown in popularity as a tool through which to help address a domain of problems where users may encounter difficulties in decision making processes. A variety of studies have shown nudges to be effective in changing user behavior [4]. In the security field, nudges have been used to help users make more secure choices. Ur et al. assess the effectiveness of visual password meters in encouraging users to create stronger passwords [103]. Turland et al. evaluate the effects of minor user interface changes on Android devices to encourage users to select more secure wireless networks [102]. In the privacy field, nudges have been used to guide users towards better privacy outcomes. Almuhimedi et al. examine the use of nudges on location disclosure decisions for mobile devices [9]. In a related study, Balebako et al. survey the extent to which nudge interventions are currently employed online and on mobile devices [11]. Each of these studies provide instances of cases where nudge interventions have been applied to guide users to better privacy and security outcomes. These studies inform the design of our nudges throughout our experiments.

While the application of nudges within privacy and security has often been 'one-size-fits-all', recent work has helped classify a variety of nudge types. Acquisti et al. review the application of nudges to problems in privacy and security and describe six categories of nudge interventions: information, presentation, defaults, incentives, reversibility, and timing [4]. This differentiation between nudge types raises the intriguing possibility that nudges may be personalized to users.

2.2.2 Individual Differences in Decision Making and Personality

Within the psychology literature, inventories have been developed to capture and quantify measures of individual differences in decision making and personality. It stands to reason that, if individuals differ based on personality and decision making traits, such differences may also translate to differential reactions to behavioral interventions across different domains of decision making - including privacy.

Personality inventories such as the well-known "Big Five" scale measure subjective personality traits which have been shown to influence behavior such as job performance [12]. In contrast, decision inventories attempt to measure aspects of users' decision making ability. We utilize both types of measures as a basis for drawing inferences regarding data disclosure decisions.

Multiple inventories have been developed within the psychology literature to measure both decision styles and decision skills. The Need for Cognition (NFC) and General Decision Making Style (GDMS) scales each measure aspects of individual decision making style [19, 88]. While NFC measures users along a single dimension, GDMS captures multiple attributes that describe decision making style.

Decision skill inventories vary from those which measure decision style in that they attempt to capture skills that individuals possess relevant to decision making. One measure of decision skills is the Adult Decision-Making Competence (A-DMC) inventory created by Bruine de Bruin et al. [18]. This inventory measures competence across seven decision skills, assessing competence in a way which the authors show is predictive of decision outcomes. Another common decision skill measure is numeracy [60]. This skill has been shown within the risk communication literature to be associated with risk perception [84, 51]. Users' perceptions of risk may have direct implications for how they approach privacy decisions.

Several of these inventories have been applied to privacy research, where Egelman et al. draw a link between privacy preferences and individual differences in decision making and personality [30]. In addition to identifying decision traits that are predictive of privacy preferences, the authors note the potential of individual differences as a tool to create personalized behavioral interventions for privacy and security decisions.

2.2.3 Tailored Nudges

In the past few years, studies have explored differential effects of nudges applied to decision problems in a number of domains. Allcot et al. assess the welfare impacts of nudges used within home energy conservation reports [8]. By capturing individual willingness to pay for the home energy reports, along with the magnitude of potential welfare gains, the authors design an algorithm to best target the nudges. In a similar study, Beshears et al. examine nudges for retirement savings plans and find the nudges to be more effective for certain demographics [14].

Within security, several studies have explored applications of tailored nudges that incorporate information about users' psychometric traits. Malkin et al. tests nudges for browser warnings that are personalized to the traits of the GDMS [66]. Although the authors identified several significant correlations between their nudges and the GDMS, they failed to find significant evidence of tailoring.

A later study by Peer et al. examined tailoring for password nudges [82]. In this study, the authors measure participants along multiple psychometric scales and were successful in identifying effects of tailoring for a subset of these. We employ several of the same measures used by the authors within our studies.

Existing work within the privacy literature has examined tailored nudges applied to disclosure decisions. Knijnenburg et al. identified differential effects for different 'justifications' within the context of disclosure rates for data on a hypothetical mobile app [52]. In practice, several of these justifications acted in a similar fashion to nudges. The authors found that characteristics such as gender moderated the effectiveness of some justifications. These findings further point to the potential for privacy nudges to be tailored to the individual traits of users. In a separate study, Coventry et al. examined the effects of a single nudge on cookie acceptance for web browsers [23]. The authors measure several personality traits, but fail to find evidence that the traits moderate the effectiveness of the nudge. We expand upon these works by focusing on multiple different psychometric traits as a basis for tailoring privacy nudges. Within the context of privacy, we are one of the first to explore the application of nudges tailored to individual differences in decision making and personality.

2.3 Study Design Overview

Our examination of tailored privacy nudges takes place within the context of data disclosure decisions. To facilitate this, each of our studies begins by presenting participants with a hypothetical or real disclosure choice and asking them to make a 'Yes' or 'No' decision on whether or not they wish to disclose their data. We insert nudges into these scenarios by modifying the choice text to elicit differences in disclosure rates between groups of participants. We employ two types of nudges across our three studies: 'framing' nudges and 'social norms' nudges. The first of these modifies the choice text to leverage framing effects, while the second introduces additional information to the choice to establish social norms which might affect participants' disclosure behavior. In Study 1, we test only the framing nudge. In Study 2 and Study 3, we test both the framing and social norms nudges.

For each nudge, we create two different wordings or 'variants' of the nudge text. Relative to each other, these variants create the desired psychological effects we wish to leverage within our nudges. For example, we create the framing nudge for Study 1 by asking one group of participants whether they wish to 'Opt-In' to a service while asking the other group whether they wish to 'Opt-Out'. The framing effect exists only in the contrast between the two wordings. We apply a similar method to create the social norms nudge used in Study 2 and Study 3. We treat each nudge 'variant' as an experimental condition within our study. Participants are randomly assigned to a single nudge variant in a between-subjects design. For Study 1, this translates into two experimental conditions across one nudge. For Study 2 and Study 3, this translates into four experimental conditions across two nudges. Across the three studies, participants are only assigned to one disclosure decision that contains a single nudge variant. We provide additional details on the construction of the nudges and the nudge conditions in the study descriptions below.

Following the decision task, we ask all participants questions from psychometric inventories designed to measure their decision making and personality traits. We selected inventories for our studies that either relate to the cognitive effect being leveraged in the nudge or have been examined in other studies exploring psychometrically targeted nudges. In Study 1, we measure two scales from the A-DMC: Resistance to

Framing and Applying Decision Rules [18]. For Study 2 and Study 3, we measure Resistance to Framing and Recognition of Social Norms (from the A-DMC) along with scales to capture Scientific Reasoning [29], Need for Cognition [19], Numeracy [60], General Decision Making Style [89], and the Big Five personality traits [12]. Additional details on each of these measures are provided in the study descriptions below. After answering questions from the psychometric inventories, participants conclude the study by answering questions on demographics and privacy attitudes.

2.4 Study 1

With Study 1, our goal was to identify whether psychometric variables could be used to infer differences in disclosure rates for a single nudge. Because tailoring privacy nudges implies the ability to select one nudge from multiple options, we planned to conduct further experimentation with multiple nudges in later studies. We do this in Study 2 and Study 3.

2.4.1 Study Design

Disclosure Scenario

We built the disclosure scenario for Study 1 around a hypothetical IoT service called 'Auto-Checkout'. This hypothetical IoT scenario was developed over several iterations and consisted of an automated checkout service that would allow users to bypass checkout lines in exchange for their indoor location information. Participants were given details of the potential benefits and drawbacks of the service, along with the choice to enroll or not with 'Yes' and 'No' response options. The details of the scenario along with the choice were presented to users on an interactive phone screen mockup.[1] While the service would allow for faster trips to the grocery store, participants were informed that their data would be used to enable personalized advertising. [2]

Nudge Design

Within the choice text of the scenario described above, we embedded a nudge designed to leverage framing effects [50]. These manipulations are frequently used within the social science literature and typically yield strong effects [28]. Because a disclosure choice does not contain a natural 'baseline' condition against which to measure the effectiveness of our nudge, we create two 'variants' of our nudge ('Opt-In' and 'Opt-Out') with which we could measure differences in disclosure rates. We use the 'Opt-In' and 'Opt-Out' variants of the nudge as our experimental conditions. When participants were presented with the scenario, they were randomly assigned to receive either the 'Opt-In' or 'Opt-Out' variant of the choice text. These nudges manipulate the text of the disclosure question to change the 'default' state of disclosure from the perspective of the user. The text for both the 'Opt-In' and 'Opt-Out' conditions is given below.

> "Do you wish to **OPT-IN** to Auto-Checkout?"
>
> "Do you wish to **OPT-OUT** of Auto-Checkout?"

Psychometric Variable Selection

Following the disclosure choice, participants were presented with questions designed to measure different psychometric traits. We selected the Resistance to Framing and

[1] Survey participants were provided with instructions on how to use the phone screen mock-ups prior to completing the disclosure task. These instructions and question format were tested in a pilot study prior to the experiment.

[2] The mock-ups, along with the full text of the disclosure scenario, are available in Appendix A.1.1.

Applying Decision Rules scales from the A-DMC for the purposes of Study 1 [18, 17]. While the full A-DMC inventory contains questions to measure individual decision making ability across seven psychometric variables, the two we selected most closely measure participants' ability to recognize the cognitive bias leveraged by our framing nudge. These scales were also shown by the authors to be highly intercorrelated. When applied to a framing nudge, we might expect the nudge to have less of an effect on participants with higher Resistance to Framing and Applying Decision Rules scores compared to participants with lower scores. For the Resistance to Framing scale, the questions are separated into two identical sets which vary only in the framing of the text. These sets must be separated by an intermediary task or time interval. Because of this, in addition to a desire to reduce potential participant fatigue, we split our survey into two sections.[3]

To protect against participants randomly selecting answers for the survey questions, we designed attention check questions which we placed within the resistance to framing questions in both the initial and followup surveys. These checks are written in the style of the surrounding questions, but contain a strictly dominated choice. Participants that are paying attention should select this answer. By using these attention checks, we can filter out participants that do not honestly complete the survey (amounting to random noise within the data) while checking for comprehension of the task.

We collected basic demographic information from participants to serve as additional controls in our regressions analysis. Of potential interest are age and level of education. Either of these variables may capture a higher level of technological literacy which may help explain disclosure choice [80]. To capture level of education, participants are asked to list their highest attained degree. These responses are

[3] Sample questions from the Resistance to Framing scale are available in Appendix A.2.1.

translated into years of education within the analysis to better facilitate regression. This translation is conducted by assigning year equivalents to each level of education between 12 and 20 years.

We additionally measured participants' scores along the Concern for Information Privacy (CFIP) scale to capture general attitudes regarding data collection and storage [92]. This variable is used as a control within our regression models.

Survey Structure

The survey for Study 1 was administered in two parts. In the initial survey, participants were presented with a hypothetical IoT disclosure choice followed by the psychometric variable scales for Resistance to Framing and Applying Decision Rules. The initial survey concluded with demographic questions.

After several days, participants were invited back to complete the followup survey in which participants answered additional questions from the Resistance to Framing scale and questions from the CFIP scale. We placed the IoT disclosure question at the front of the initial survey and the CFIP questions at the end of the followup. By presenting the questions in this order, we avoided priming participants with context which may have influenced their disclosure choice.

2.4.2 Results

Sample Demographics

We recruited a sample of 200 participants from Mechanical Turk to complete the initial survey. Although samples drawn from Mechanical Turk are not representative, numerous behavioral effects (including framing effects) have been replicated on the platform [78]. Of the participants recruited into the study, 176 completed the initial survey without incorrectly answering any of the attention check questions. Several days after the initial survey, this subset of participants was invited back to complete the followup survey. The response rate to this second stage survey was 81%, yielding a sample of 145 participants. Of these, only 2 participants failed to answer the attention check questions correctly. After excluding these participants, our analysis was conducted on a final sample of 143 participants.[4] Of this final sample, 60% (86) were male and the median age was 34. All participants had at least a high school degree, with 48% (69) having achieved a bachelors degree or higher.

Summary Statistics

Across the two experimental conditions, 55% (78) of participants chose to allow the data disclosure within the hypothetical IoT scenario. We measured the size of the framing effect by comparing the likelihood of participants to allow the data disclosure between nudge conditions. Of the 143 participants considered in our analysis, 64% (46) of those assigned to the 'Opt-In' condition chose to allow the disclosure while only 44% (32) assigned to the 'Opt-Out' condition chose to do the same. The difference in disclosure rates between these conditions is significant with a p-value of $p = 0.02291$ when tested using Chi-squared, suggesting a significant main effect of the nudge.

We additionally measured the means and standard deviations for the two psychometric variables and the privacy preferences scale. The mean score for Resistance to Framing within the sample was 3.865 with a standard deviation of 0.403. The possible range of scores on this scale is 1 to 6. This result suggests a sample which is slightly resistant to framing effects overall. For the Applying Decision Rules scores, the sample mean was 0.615 with a standard deviation of 0.246. The range of possible

[4] Our analysis for each study was conducted in R using publicly available packages for regression analysis and data presentation [35, 36, 46, 110, 108].

scores for this variable is 0 to 1 where higher scores indicate greater ability to apply rules to decision problems. The sample mean for privacy attitudes as measured by the CFIP scale was 5.887 with a standard deviation of 0.984. The range of possible scores for this scale is 1 to 7. These scores potentially indicate a higher privacy sensitivity within the sample.

Logistic Regression

We used logistic regression models to conduct a difference-in-differences analysis of the response data. In each regression, the coefficient of interest to our research question was the interaction between the psychometric variable scores and the nudge condition assigned to the participant. This allowed us to see whether the effect of the assigned nudge on disclosure likelihood varied significantly with the psychometric variables.

In total, we examined two sets of models – one for each psychometric variable (Resistance to Framing and Applying Decision Rules). For each of the psychometric variables, we examined four regressions models, varying the number of control variables in each one. In the first model, we assessed the interaction between the framing nudge and psychometric variable score alone. In the subsequent models, we incrementally introduced controls for privacy attitudes (CFIP) and demographics (age, gender, and education). We represent the assigned experimental condition (whether participants were shown the 'Opt-In' or 'Opt-Out' variant of the nudge) in our models as a binary coded dummy variable (Opt-In Condition).

We used Generalized Additive Models (GAMs) to test for omitted transformations on the continuous regressors. This analysis revealed a non-linear transformation on the ADR psychometric variable. To best satisfy the 'linearity in parameters' assumption of the logistic regression model, this score variable was log transformed. A subsequent test showed the Log ADR score to be more linear. We use the Log ADR score throughout the remainder of our analysis.

Table 2.1 presents the regression coefficients expressed in log-odds along with standard errors for the models containing all controls for each psychometric variable.[5] Each column of the table presents the regression model for a separate psychometric variable. Overall, we did not identify significant effects for either of the interaction terms in our models.

Table 2.1: Logistic Regression Coefficients for Study 1

	Dependent variable:	
	disclosure	
	(1)	(2)
Opt-In Condition (OC)	3.909	0.745
	(3.496)	(0.560)
Framing Score (FR)	0.266	
	(0.607)	
Log-ADR Score (ADR)		1.001*
		(0.565)
CFIP Score	−0.238	−0.143
	(0.192)	(0.190)
Age	0.007	0.006
	(0.019)	(0.019)
Female	0.391	0.493
	(0.378)	(0.383)
Years of Education	0.141	0.076
	(0.093)	(0.096)
AC*FR	−0.768	
	(0.898)	
AC*ADR		−0.250
		(0.787)
Constant	−2.352	−0.381
	(2.835)	(1.818)
Observations	143	143
Log Likelihood	−92.626	−90.389
Akaike Inf. Crit.	201.252	196.778
Note:	*p<0.1; **p<0.05; ***p<0.01	

While none of the interaction coefficients we examined were significant, the direc-

[5] Regression models with incremental levels of controls for Study 1 are available in Appendix A.4.1. These models did not yield significant effects for the interaction terms.

tion of the coefficients make intuitive sense. For both the Resistance to Framing and Log-ADR variable scores, the sign of the interaction coefficient is negative, indicating that a higher decision competence for these variables translates into a decrease in the effectiveness of the framing nudge on disclosure behavior.

Although we did not identify significant effects for the interactions between the two psychometric variables and nudge conditions, our results may have been limited by factors including our sample size and the hypothetical nature of the disclosure scenario. Additionally, limitations within our nudge design and selection of psychometric variables may have impacted our results. We attempt to address these potential limitations and build upon the design of Study 1 within Study 2 and Study 3.

2.5 Study 2

Although Study 1 did not yield significant findings, the direction of our regression coefficients showed promise. In Study 2, we sought to determine whether the null result observed in Study 1 was robust by focusing on the potential limitations of Study 1 that may have contributed to our null results. These potential limitations include the design of our nudges and disclosure choices along with our selection of psychometric variables. If we selected the wrong nudges or psychometric variables, this may have lead to a false negative. In Study 2, we attempt to minimize these factors.

We addressed these issues in part by switching from hypothetical to real disclosure choices and including an additional nudge within our study design. We additionally expanded upon the selection of psychometric variables to capture a broader range of psychometric traits. Because of the high risk of identifying spurious correlations in our analysis within Study 2, we made the decision prior to collecting data to follow up Study 2 with a subsequent replication study (Study 3), which would attempt to confirm any effects possibly identified in Study 2.

2.5.1 Study Design

Disclosure Scenario

Our use of a hypothetical disclosure scenario in Study 1 potentially limited the size of the effect of our framing nudge. Although we iterated on the design of the scenario to make it more realistic, the hypothetical scenario lacked the perception of risk that users might face when making real disclosure decisions. Research in the privacy literature indicates that this may impact how users make privacy choices [75][6]. Users may respond differently to a real disclosure decision than they would to a hypothetical one. In the context of our study design, asking participants to make a real disclosure decision may yield stronger main effects for our nudges compared to a hypothetical decision. For Study 2, we employ deception to construct a scenario and disclosure question with real perceived risks by participants.

The design of our disclosure question and experimental manipulation for Study 2 is based off the design used by Samat et al. in a study of framing effects under different levels of risk [85]. In that study, the authors asked participants to answer a series of 'ethical behavior questions' drawn from Acquisti et al. [2]. Those questions asked participants how frequently they engaged in a variety of behaviors of varying degrees of sensitivity such as "Have you ever had a one-night stand". Before answering the questions, participants were presented with the ostensible choice to decide whether or not they wished to share their responses to the ethical behavior questions with a third-party. Samat et al. manipulated the identity of the third-party to raise or lower the perceived risk of the disclosure to participants. In the high-risk condition, participants were told that their responses would be shared with other users on Mechanical Turk community forums. Using this design, the authors found that the high-risk condition yielded the largest framing effect. No matter participants' answers to the disclosure choice, their responses were not actually shared.

In an early iteration of our design for Study 2, we combined the high-risk disclosure condition and framing nudge of Samat et al. with an expanded selection of psychometric variables. A pilot of this design on a small sample revealed that participants were skeptical of the deception. Several participants questioned why researchers would want to share their responses to ethical behavior questions with other users on Mechanical Turk community forums. From this feedback, we changed the third-party within the scenario from users of Mechanical Turk community forums to researches at other universities. A test of this deception yielded better results. Our final version of the disclosure scenario for Study 2 includes this change along with additional text designed to increase the perceived sense of risk for disclosure.

Nudge Design

We created two nudges to embed within the disclosure choice of our deceptive scenario. As in Study 1, these nudges each contain two variants that enable us to measure a relative difference in disclosure rates. This yielded a total of four experimental conditions spread across two nudges in a between-subjects design. Because of its strong effects within the literature and in Study 1, we again used a framing nudge within Study 2. The design of our framing nudge mimics that used by Samat et al. and is shown below.

> "**Allow** your responses to the ethical behavior questions to be shared with

> researchers outside of our study team, along with your Mechanical Turk ID?"
>
> "**Prohibit** your responses to the ethical behavior questions from being shared with researchers outside of our study team, along with your Mechanical Turk ID?"

The above 'Allow' and 'Prohibit' variants differ only in their presentation of the disclosure choice. Participants assigned to one of these conditions were then presented with the response options 'Yes' and 'No'. We randomized the order of these response options to avoid any potential ordering effects.

Research from the psychology literature on social norms has shown them to have some of the strongest effects among different nudges [96]. We constructed a second nudge that leverages those effects to create differences in disclosure rates similar to the framing nudge. This nudge consists of 'High Norms' and 'Low Norms' conditions that establish a social norm regarding the percentage of other participants that chose to disclose their responses. Unlike the framing nudge, the text for this nudge is appended to the end of the description of the disclosure scenario. We tested several versions of this nudge before arriving at a version that consistently yielded strong effects. The text of the experimental manipulations is shown below.

> "In our past studies, **73**% of participants chose to allow their responses to be shared with researchers outside of our study team."
>
> "In our past studies, **31**% of participants chose to allow their responses to be shared with researchers outside of our study team."

Because the disclosure question no longer contained the experimental manipulation, both the question and the response options for the social norms nudge conditions remained fixed. As for the conditions associated with the framing nudge, we randomized the order of the response options to avoid potential ordering effects. The full text of each experimental condition – including the text of the disclosure scenario and manipulations – is provided in Appendix A.1.2.

Psychometric Variable Selection

We expanded our selection of psychometric variables in Study 2 to cover a broader range of cognitive traits than in Study 1. In total, we selected 15 psychometric variables from the decision science and psychology literatures that measure a variety of decision making and personality characteristics. This blend of variables used scales that capture both objective skill measures (such as the scales of the A-DMC) and self-reported personality measures.

Due to its direct applicability to the framing nudge, we retained the Resistance to Framing scale that we used in Study 1. To this we added scales to measure Recognition of Social Norms (SN) [18], Scientific Reasoning (SRS) [29], Need for Cognition (NFC) [19], Numeracy (NUM) [60], General Decision Making Style (GDMS) [89], and the Big Five personality traits (B5) [12]. These additional scales were included due to either their relevance to the nudge being tested or their previous use in related studies on personalized nudges in security [66][82]. In the case of the Recognition of Social Norms scale, we expected these scores to have a strong relationship with the social norms nudge – where higher scores translate to greater effectiveness of the nudge on impacting disclosure rates.

Each of the scales with the exception of Numeracy consisted of multiple choice questions. The Big Five measures five personality traits that we consider independently. These are extraversion, agreeability, conscientiousness, neuroticism, and openness. Likewise, the GDMS identifies five decision making styles. These are rational, intuitive, dependent, avoidant, and spontaneous. We also consider these independently within our analysis. Both the resistance to framing and recognition of social norms scales are split into two sets of related questions that were administered to participants at different times.[6] In addition to the psychometric variables, we collect demographic information and participants' scores along the Internet Users' Information Privacy Concerns (IUIPC) scale [65] to use as controls within our regression models. We selected the IUIPC scale over the CFIP scale from Study 1 due to its wider use within the privacy literature.

Survey Structure

As in Study 1, the survey was administered in two parts to reduce the cognitive burden on participants and to allow for time separation as required by some of the psychometric variable measures. In the initial survey, participants completed the disclosure choice, ethical behavior questions, need for cognition scale, numeracy scale, and scientific reasoning scale. Participants also completed the initial sections of the scales designed to measure resistance to framing effects and recognition of social norms. The following day, participants were invited back to complete a followup survey which contained the second halves of the resistance to framing and recognition of social norms scales. In addition, the followup survey contained the measures for GDMS, Big Five, and IUIPC.

[6] Sample questions from the Recognition of Social Norms scale are available in Appendix A.2.2

2.5.2 Results

Sample Demographics

We recruited 1,200 participants to complete Study 2 from Mechanical Turk. Of those recruited, 1,198 completed the initial survey and 966 completed the followup survey and passed the minimum threshold for attention check questions.[7] Of the 1,198 that completed the initial survey, 50% were male and the median age was 35. 58% (700) had completed a Bachelor's degree or higher. To facilitate our exploratory analysis, we split the sample by survey. Because the psychometric variable scales were split between the two surveys (and because the experimental manipulation was contained in the initial survey), several psychometric variables can be examined without needing data from the followup survey. By splitting our sample into a 'complete' sample (all 1,198 participants) and a 'partial' sample (the 966 participants that completed both surveys), we can take advantage of the larger sample size for a subset of the psychometric variables.

Summary Statistics

For both the framing and social norms nudges, we observed significant differences in the disclosure rates between conditions on the complete sample. Of the 597 participants that were assigned to the framing nudge, 66% (197) of those assigned to the 'allow' condition chose to disclose their responses while 39% (115) of those assigned to the 'prohibit' condition chose to disclose their responses. For the 601 assigned to

[7] We developed 8 attention check questions for Study 2 that we placed throughout our survey. These attention check questions were designed to measure task comprehension and presented participants with a task in the style of surrounding questions that contained a strictly dominated answer. We used our attention check questions as a robustness check on the results of our analysis. While we present the results of participants that answered at least 1 attention check question correctly (which excludes the fewest participants), the results are consistent for higher numbers of attention check questions.

the social norms nudge, 65% (197) of those assigned to the 'high norms' condition chose to disclose their responses while 47% (140) of those assigned to the 'low norms' condition chose to disclose their responses. Both differences in disclosure rates are significant when tested using Chi-squared. The differences between the conditions of the framing and social norms nudges are significant with p-values of $p = 7.584e - 11$ and $p = 5.199e - 06$ respectively. We observed similar difference in disclosure rates and significance levels for the partial sample.

Table 2.2 shows summary statistics for the psychometric variables and privacy concerns measured in Study 2 – including the mean, standard deviations, and ranges of possible scores for each variable. Higher scores for these variables translate to higher competence or greater intensity for the corresponding decision making and personality traits.

Table 2.2: Psychometric Variable Summary Statistics for Study 2

Variable	Mean	St. Dev.	Range
Scientific Reasoning	6.381	2.718	0 – 11
Need for Cognition	4.611	1.250	1 – 7
Numeracy	8.171	2.033	0 – 11
Resistance to Framing	4.924	0.486	1 – 6
Recognition of Social Norms	0.471	0.271	-1 – 1
GDMS Rational	3.888	0.674	1 – 5
GDMS Intuitive	3.293	0.904	1 – 5
GDMS Dependent	3.192	0.855	1 – 5
GDMS Avoidant	2.479	1.060	1 – 5
GDMS Spontaneous	2.510	0.890	1 – 5
Big 5 Extraversion	2.893	0.946	1 – 5
Big 5 Agreeableness	3.710	0.755	1 – 5
Big 5 Conscientiousness	3.918	0.771	1 – 5
Big 5 Neuroticism	2.668	0.970	1 – 5
Big 5 Openness	3.717	0.725	1 – 5
IUIPC	5.879	0.911	1 – 7

We computed the correlation strength and significance for each pairing of psy-

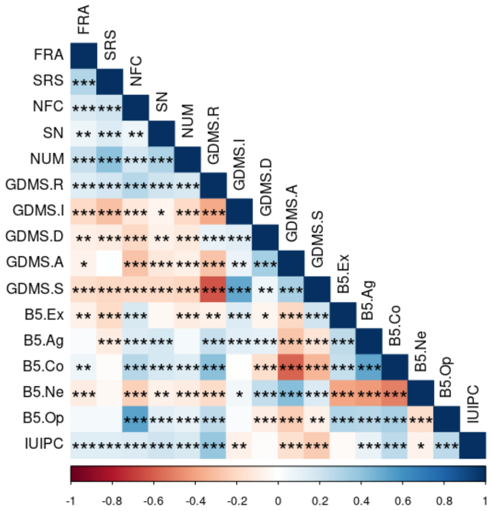


Figure 2.1: Correlation Strength and Significance for Psychometric Variables
*Note:**p<0.1; **p<0.05; ***p<0.01

chometric variables to assess the coverage of cognitive traits within our study. This correlation table is displayed in Figure 2.1. With the exception of two psychometric variables within the GDMS and Big Five, most of the variables are only weakly correlated with each other. This suggests that the selected scales are measuring a broad range of cognitive traits, with little overlap.

Logistic Regression

We tested multiple logistic regression models as part of our exploratory analysis to examine the relationship between disclosure choice, our nudges, and the measured psychometric variables. For each model, our dependent variable was participants' disclosure choice. Our explanatory variables were the assigned nudge condition and the relevant psychometric variable score. Across the two nudges and 15 psychometric variables, we constructed 30 sets of regression models. We used the IUIPC score and demographic variables as controls within our regressions. The interaction term between the assigned nudge condition and psychometric variable score is the primary coefficient of interest to our research question. As in Study 1, we represent the assigned experimental condition as a binary coded dummy variable ('Allow Condition' for models with the framing nudge and 'High Norms Condition' for models with the social norms nudge).

Out of the regression models that we created, we identified 3 pairs of psychometric variables and nudge conditions with potentially significant interaction terms. These included the Recognizing Social Norms score and social norms nudge, the Big Five Extraversion score and framing nudge, and the Big Five Conscientiousness score and framing nudge. The coefficients for these regression models expressed in log-odds along with the standard errors for each are presented in Table 2.3.[8]

Column 1 of Table 2.3 shows the regression model with control variables for the pair of the social norms nudge and Recognizing Social Norms score. Columns 2 and 3 show the regression models with controls for the framing nudge paired with the Big 5 Extraversion and Conscientiousness scores respectively.

Of the potentially significant effects, the interaction between the Recognizing Social Norms score and the social norms nudge is the most intuitive and has the strongest effect. The direction of the coefficient implies that as a participant is more likely to recognize social norms, they may be more likely to be influenced by a social norms nudge. While less intuitive, the effects between the two Big Five traits and the framing nudge can also be interpreted. The direction of the regression coefficient for Big Five Extraversion score and the framing nudge condition suggests that those who are more extroverted may be more likely to be influenced by a framing nudge. Likewise,

[8] Results for the regression models using incremental levels of controls for the three potentially significant psychometric variables are available in Appendix A.4.2. Regression results for the non-significant pairs of psychometric variables and nudge conditions are available from the authors upon request.

Table 2.3: Logistic Regression Coefficients for Study 2

	Dependent variable:		
	disc		
	(1)	(2)	(3)
Recognizing Social Norms	-1.232^{**}		
	(0.574)		
Big 5 Extraversion		-0.056	
		(0.140)	
Big 5 Conscientiousness			0.258
			(0.193)
High Norms Condition	-0.322		
	(0.420)		
Allow Condition		-0.398	2.849^{***}
		(0.616)	(1.061)
IUIPC	-0.527^{***}	-0.485^{***}	-0.485^{***}
	(0.128)	(0.122)	(0.123)
Age	-0.021^{**}	-0.008	-0.007
	(0.010)	(0.009)	(0.009)
Female	-0.083	0.200	0.157
	(0.208)	(0.203)	(0.201)
African American	-0.012	0.162	0.171
	(0.363)	(0.369)	(0.368)
Hispanic	0.222	0.027	-0.052
	(0.405)	(0.365)	(0.366)
Asian	0.489	0.451	0.355
	(0.493)	(0.472)	(0.469)
Other Race	0.147	0.130	0.094
	(0.621)	(0.648)	(0.647)
High School	-12.697		
	(535.411)		
Associate Degree	-12.650	0.166	0.116
	(535.411)	(0.327)	(0.327)
Bachelor's Degree	-12.768	0.274	0.283
	(535.411)	(0.302)	(0.302)
Advanced Degree	-13.118	-0.086	-0.013
	(535.411)	(0.391)	(0.387)
Other Education		13.218	13.081
		(535.411)	(535.411)
Recognizing Social Norms*High Norms Condition	2.432^{***}		
	(0.784)		
Big 5 Extraversion*Allow Condition		0.487^{**}	
		(0.204)	
Big 5 Conscientiousness*Allow Condition			-0.464^{*}
			(0.262)
Constant	16.968	2.615^{***}	1.425
	(535.411)	(0.864)	(0.963)
Observations	453	482	482
Log Likelihood	-281.000	-301.875	-304.746
Akaike Inf. Crit.	591.999	633.749	639.492
Note:			*p<0.1; **p<0.05; ***p<0.01

the direction of the coefficient for the Big Five Conscientiousness score and framing nudge condition suggest that participants who are less conscientious may be more likely to be influenced by a framing nudge.

We employed the same statistical methods as in Study 1 to test for misspecification of the logistic regression models. We used generalized additive models to detect omitted transformations of the independent variables. This analysis indicated the each of the independent variables was roughly linear.

2.6 Study 3

Our goal in Study 2 was to determine whether the null effects we observed in Study 1 were robust or rather due to limitations in our study design. While we cannot prove the absence of an effect, we sought to minimize the likelihood of a false negative in Study 2 by expanding our selection of nudge types and psychometric variables. We found three potentially significant interactions between nudge conditions and psychometric variables. However, by testing additional nudges and psychometric scales, we risk identifying spurious correlations. Study 3 builds upon the results of Study 2 by attempting to replicate the three potentially significant effects we identified using a separate sample.

2.6.1 Study Design

For this study, we made only minimal changes to the study design and survey structure used in Study 2. Our primary change was to condense the survey questions into a single instrument by removing the scales for the psychometric variables we did not wish to test. Condensing the survey removed the need for a separate followup survey and allowed participants to complete the study in one session. In total, the survey used for Study 3 kept the same disclosure choice and nudge manipulations while measuring recognition of social norms, the Big 5 personality traits, and the IUIPC. Because many of our attention check questions from Study 2 were embedded within the scales for the other psychometric variables not considered in Study 3, we kept some of the questions from the other scales to preserve these attention checks. We additionally expanded our sample size to 2,000 participants (1,000 per nudge). This decision was informed in part by a post-hoc power analysis conducted using the observed effects sizes from Study 2.[9] We pre-registered this study design along with our analysis plan on Open Science Framework prior to collecting data.[10]

2.6.2 Results

Sample Demographics

We recruited 2,000 participants to complete our replication study using Mechanical Turk. During recruitment, we excluded participants that had previously participated in Study 1 or Study 2. Out of this total, 1,996 participants provided complete responses and passed the minimum threshold for attention check questions. The makeup of this sample was 50% male with a median age of 36. 61% $(1,222)$ had completed a Bachelor's degree or higher. Because we were able to condense our psychometric variable scales into a single survey, we did not split our sample as in Study 2.

Summary Statistics

We observed similar effect sizes on both the framing and social norms nudges compared to Study 2. For the framing nudge, 59% (293) of participants assigned to the 'allow' condition chose to share their responses to the ethical behavior questions while

[9] The results of this power analysis are available in Appendix A.3.

[10] `https://osf.io/nfyma`

only 33% (165) of those assigned to the ‘prohibit’ condition chose to do the same. For the social norms nudge, 65% (321) of those assigned to the ‘high norms’ condition chose to disclose their responses while only 47% (239) of those in the ‘low norms’ condition shared their responses. Using a Chi-Squared test, both differences were, as in Study 2, significant at $p = 6.051e - 16$ and $p = 6.432e - 8$, respectively.

We additionally examined the means and standard deviations for the Recognizing Social Norms, Big Five Extraversion, and Big Five Conscientiousness psychometric variables. For the Recognizing Social Norms variable, we observed a sample mean of 0.474 and a standard deviation of 0.294. On the Big 5 Extraversion and Conscientiousness scales, we observed sample means of 2.874 and 3.829 with standard deviations of 0.864 and 0.773 respectively. For the IUIPC scale, we observed a mean privacy concern score of 5.767 with a standard deviation of 1.033. Many of these statistics were similar to the values observed in Study 2. Using a two-sample t-test, we did not find significant differences in the means of the three psychometric variables between Study 2 and Study 3. This suggests that the scores for these three variables observed in Study 2 are consistent with those in Study 3.

Logistic Regression

We tested the same logistic regression models that we used in Study 2 for the three potentially significant pairs of psychometric variables and nudge conditions (Recognizing Social Norms score and social norms nudge, the Big Five Extraversion score and framing nudge, and the Big Five Conscientiousness score and framing nudge). Overall, we failed to replicate the effects from Study 2. For the three logistic models, the interaction terms were either not significant, weakly significant (at the $p < 0.1$ level), or lost significance when controls were added. The coefficients for these regression models expressed in log-odds along with the standard errors for each are presented in Table 2.4.[11]

Whereas the interaction term for the Recognizing Social Norms score and social norms nudge pair was the strongest of the three effects in Study 2, the same effect was only significant at the $p < 0.1$ level when control variables were excluded from the regression in Study 3. When the control variables were added, the coefficient for the interaction term was no longer significant. For the Big Five Conscientiousness score and framing nudge pair, the direction of the interaction coefficient changed from negative to positive. Again, the interaction term for this variable and condition pair was only weakly significant for the regression model without control variables. When control variables were added, this term lost significance.

Only the significance of the interaction term for the Big Five Extraversion score and framing nudge pair remained comparable from Study 2 to Study 3. However, it did so only at the $p < 0.1$ level. While the interaction term was significant at the $p < 0.05$ level when no control variables were included, the significance level diminished when they were reintroduced.

To explore these effects in more detail, we conducted a secondary exploratory analysis where we split each of the three psychometric variables into tertiles. This allowed us to compare participants who scored in the bottom third for a particular variable to those in the middle and top thirds within our regression models. By splitting the data, we can potentially detect more detailed effects within the interaction. This analysis revealed significant effects for the interaction term between the middle and bottom tertiles for the Big Five Extraversion score and framing nudge pair at the $p < 0.01$ level. Although this is potentially highly significant, it only encompasses a portion of the range of potential Big Five Extraversion scores and indicates that the

[11] Results for the regression models using incremental levels of controls are available in Appendix A.4.3.

Table 2.4: Logistic Regression Coefficients for Study 3

	Dependent variable:		
	disc		
	(1)	(2)	(3)
Recognizing Social Norms	-1.054^{***}		
	(0.377)		
Big 5 Extraversion		-0.062	
		(0.117)	
Big 5 Conscientiousness			-0.141
			(0.130)
High Norms Condition	0.363		
	(0.300)		
Allow Condition		0.283	0.364
		(0.475)	(0.692)
IUIPC	-0.953^{***}	-0.500^{***}	-0.496^{***}
	(0.101)	(0.071)	(0.072)
Age	0.011^{*}	-0.010	-0.009
	(0.006)	(0.006)	(0.006)
Female	-0.303^{**}	-0.016	-0.015
	(0.148)	(0.140)	(0.140)
Non-Binary	0.859	1.206	1.273
	(1.250)	(1.223)	(1.202)
African American	-0.414	-0.350	-0.279
	(0.265)	(0.237)	(0.233)
Hispanic	0.096	-0.107	-0.091
	(0.320)	(0.285)	(0.284)
Asian	-0.825^{**}	-0.088	-0.084
	(0.325)	(0.271)	(0.271)
Other Race	0.144	-0.476	-0.472
	(0.490)	(0.669)	(0.671)
High School	0.590	1.509	1.396
	(1.038)	(0.985)	(0.982)
Associate Degree	0.513	1.826^{*}	1.707^{*}
	(1.020)	(0.974)	(0.971)
Bachelor's Degree	0.374	1.354	1.267
	(1.016)	(0.969)	(0.966)
Advanced Degree	0.518	1.063	0.967
	(1.029)	(0.979)	(0.976)
Recognizing Social Norms*High Norms Condition	0.818		
	(0.534)		
Big 5 Extraversion*Allow Condition		0.287^{*}	
		(0.159)	
Big 5 Conscientiousness*Allow Condition			0.194
			(0.177)
Constant	5.367^{***}	1.331	1.702
	(1.160)	(1.093)	(1.118)
Observations	936	992	992
Log Likelihood	-551.202	-607.651	-609.227
Akaike Inf. Crit.	1,134.404	1,247.302	1,250.455
Note:	*p<0.1; **p<0.05; ***p<0.01		

effect may be fragile. [12]

Johnson-Neyman Analysis

In addition to our regression analysis, we examined our data for Study 2 and Study 3 using the Johnson-Neyman technique. This technique has been previously used in the literature to examine the effectiveness of tailored nudges on security behavior [82]. This method splits apart interaction terms for linear models to help identify "regions of significance" within moderator variables [43]. Applied to our study, the Johnson-Neyman technique can help identify ranges within the psychometric variable scores where the interaction with the nudge condition is significant at the $p < 0.05$ level. While this method is potentially useful for identifying effects overlooked by traditional regression models, splitting apart the interaction term in this way increases the risk of identifying spurious correlations. When applied to the three potentially significant effects identified in Study 2, the Johnson-Neyman technique did identify regions of significance for all three. However, and importantly, we were unable to replicate these results for Study 3. In this replication, the regions identified by the Johnson-Neyman technique in Study 2 were either different from or non-existent in Study 3. These results are consistent with the results of our logistic regression analysis and indicate that, if present, the effects of psychometrically tailored privacy nudges are fragile.

2.7 Discussion

Overall, the results of our three studies indicate that effects for tailored privacy nudges are difficult to identify with consistency. Although Study 2 identified three potentially significant effects, our replication of these effects in Study 3 found them to be either

[12] Regression results for the tertiles analysis are available from the authors upon request.

fragile or non-existent. This result suggests that tailored privacy nudges at the scale of our studies may not be practical in application.

We should note that the results observed in our three studies do not prove the lack of an effect for tailored privacy nudges. It is possible that a study with stronger main effects for the nudges, different psychometric variables, or larger sample sizes may find evidence for effects of psychometric tailoring on privacy nudges. Each of our study designs attempted to optimize these three dimensions to minimize the likelihood of a false negative. We constructed our nudges based on two of the strongest cognitive biases identified within the nudging literature [96]. Likewise, our selection of psychometric variables reflected a broad scope of cognitive traits and reflected the variables used by similar studies on tailored security nudges [66][82]. Our selected sample sizes were informed by power analyses using estimated effects sizes.

It is also possible that different results may be observed with a different sample population. Stewart et al. find the population of participants on Mechanical Turk available for behavioral research at any time to be relatively small (less than 10,000) [95]. Combined with a slow rate of turnover, this suggests that participants may become habituated to behavioral research over time. Participants that are repeatedly exposed to questions from the same psychometric scales may come to learn the 'correct' answers (particularly for scales that measure decision making ability). Peer et al. examine several alternative crowdsourcing platforms to Mechanical Turk, finding them to be less conditioned to behavioral research [81]. For Study 1, this conditioning effect may have contributed to the low variance we observed in scores for psychometric variables such as resistance to framing. In this way, samples drawn from crowdsourcing platforms may be different from representative samples of the general population.

While our studies employed large sample sizes, it is still possible that effects for tailored nudging could be identified using larger samples. However, as the sample sizes required to detect the effects from psychometric variables on nudges increase, tailored nudges become impractical from the perspective of many applications. The combination of large sample requirements and potentially fragile effects may make tailored nudges feasible only to organizations with vast amounts of user data such as Facebook or Google.

The nature of our results further emphasizes the importance of replication studies in behavioral research. When testing for interaction effects between multiple moderating variables, the risk increases of identifying spurious correlations with high significance. Future research should take care to validate such results.

2.8 Conclusion

We conducted three online studies with the goal of identifying effects that would indicate whether privacy nudges could be tailored according to individual differences in decision making and personality. In Study 1, we tested a hypothetical disclosure choice embedded with a framing nudge. Study 2 cast a wider net by testing more types of nudges and a wider variety of psychometric variables. This study yielded three potentially significant effects which we attempted to replicate in Study 3. Our replication did not confirm these effects, finding them to be fragile at best. Together, the results of these studies suggest that tailored privacy nudges are likely not feasible for many small-scale applications of nudges - such as those seen in the privacy literature - and reaffirms the importance of replicating potentially significant results. While on statistical grounds a null effect cannot be used to rule out their existence, the potential sample sizes required to identify robust means of tailoring privacy nudges likely make them impractical to all but the largest organizations.

Chapter 3

Explanations for Mobile Privacy Recommendations

Abstract

When faced with recommendations produced by automated systems, explanations can be an effective tool to help users evaluate different aspects of the recommendation. In the context of privacy decisions, explanations can potentially help users better determine whether or not to accept a recommendation, which may result in better privacy outcomes. While recommendations and explanations have been broadly studied across many fields, relatively little work has focused on the application of explanations to privacy recommendations. We examine explanations in the context of hypothetical permission recommendations produced by a mobile privacy assistant. In particular, we design and test three explanations of differing styles for a hypothetical permission recommendation produced for a fictional video conferencing app. We examine how "social," "behavioral," and "hybrid" explanations influence adoption rates for the hypothetical recommendation. Our results do not find strong evidence linking the three explanations to hypothetical intent. We discuss the limitations of hypothetical studies on privacy behavior and make suggestions for future work examining explanations in the context of real user behavior.

3.1 Introduction

Automated recommendations have grown in popularity as a means to help users navigate decisions where high search costs, information asymmetries, or other cognitive hurdles may prevent them from reaching optimal outcomes according to their preferences [48]. Recommendation systems have been developed to assist users in a wide variety of contexts, and often employ models of user behavior and preferences to generate recommendations [55]. More recently, research has sought to apply automated recommendations to privacy decisions, where users face unique challenges when making decisions about information disclosure [5]. Privacy recommendations may attempt to assist users with data disclosure decisions by weighing risks and benefits of disclosure which may be unknown to users with their privacy preferences [52]. In this manner, a privacy recommendation can suggest a user either 'allow' or 'deny' a particular disclosure. While different types of recommendations may address many of the same cognitive hurdles, recommendations for privacy decisions must also take into account privacy preferences which may be contextual or incomplete [73, 74]. Several existing privacy recommendation systems have sought to model privacy preferences using machine learning and studies of real user behavior [61, 62].

The growing complexity of many of the behavioral and preference models underpinning automated recommendation systems has prompted interest in providing explanations for their outputs [40, 10]. Work on explanations for automated recommendations has often sought to improve transparency in recommendations by explaining how they are generated [77, 71]. This may, in turn, increase the perceived quality of the recommendations and boost adoption by users. Existing studies on explanations have often sought to apply them to recommendations in different contexts with this goal in mind [67, 53]. Despite growing interest in explanations for recommendations broadly, relatively little work has examined the application of explanations to privacy recommendations. Furthermore, studies have yet to examine the potential effects of explanations when applied to the context of mobile privacy recommendations.

We build upon the work of Liu et al. to study the impact of explanations for mobile privacy recommendations on the hypothetical adoption intents of users [61]. Our study focuses on three 'styles' of explanations ('social', 'behavioral', and 'hybrid') and measures their impact on the adoption intents of users for mobile permission recommendations concerning a hypothetical video conferencing app called 'VideoMeet'. Our selection of explanations and disclosure context allow us to address three questions currently unexplored within the literature:

RQ1 Does the presence of explanations increase the adoption intent for the hypothetical privacy recommendation among participants?

RQ2 How do the social and behavioral explanations perform relative to each other in terms of adoption intent?

RQ3 How does the hybrid style explanation (which combines the social and behavioral explanations) perform relative to the individual explanations?

To address these research questions, we present the hypothetical disclosure scenario to participants and measure their adoption intent through an online survey. We randomly assign each participant to one of four experimental conditions where we vary the style of explanation presented alongside the recommendation (including a control condition where participants receive no explanation). Following this disclosure choice, we ask participants followup questions designed to measure their perceptions regarding the quality of the recommendation and (where applicable) explanation. We additionally ask questions that measure their perceived trust in the hypothetical recommendation system along with the existing preferences for video conferencing apps and their general preferences regarding online privacy. We use these perceptions and preferences as controls in our regression models.

We construct two logistic regression models to examine the performance of the explanation on adoption intent. In Model 1, we examine the impact of our explanations relative to no explanations. In Model 2, we examine the effectiveness of the 'hybrid' explanation relative to the 'social' and 'behavioral' explanations individually. The results of our study fail to find significant effects for our explanations on the adoption intents for our recommendations across both models. We find instead that perceptions of recommendation quality and overall privacy preferences were strong predictors of adoption intent across both the 'allow' and 'deny' recommendations. A followup analysis using both qualitative and multiple-choice response data revealed that many participants did not believe that it was necessary for the hypothetical 'VideoMeet' app to access their location data. These results may suggest a limit to effectiveness of explanations in the face of strong pre-existing preferences. Future studies on explanations for privacy recommendations may want to consider focusing on real disclosure behaviors where possible.

3.2 Related Literature

Our study builds upon and contributes to ongoing research in the fields of recommendation systems, explanations, and privacy decision making.

3.2.1 Recommendation Systems

Multiple studies have aimed to develop systems that can provide automated recommendations for different choice problems. Adomavicius et al. provide an overview of different types of recommendation systems and propose categories of recommendations based on the methods used to create the recommendations [7]. The authors propose three categories of recommendations including content-based recommendations, collaborative filtering recommendations, and hybrid approach recommendations. These incorporate data on user preferences, the preferences of similar users, and data that combines both the preferences of the user and other users respectively.

Recent studies have applied recommendation systems to different online choice situations. Jin et al. develop recommendations that incorporate personal characteristics to make music recommendations [49]. In a related study, Zhu et al. develop a system to recommend mobile applications that takes into account user security concerns [112]. In a final example, Chen et al. propose two recommendation systems for e-commerce that account for seller profitability [22].

Recommendations for Privacy Decisions

Building upon the recommendation systems literature, privacy researchers have studied the application of recommendation systems to disclosure decisions. In an early study, Xie et al. examine the privacy preferences of users around location sharing [111]. The authors identify contextual predictors of location sharing preferences and use these to develop a recommendation algorithm which has the potential to reduce users' privacy risk. In a related study, Knijnenburg et al. examine user preferences for fine-grained location sharing [52]. Across two experiments, the authors find that users' preferences are highly contextual and develop recommendations for sharing actions based on their evaluations of different scenarios.

More recently, privacy researchers have applied recommendation systems to problems of mobile privacy. In a 2016 study, Liu R. et al. develop a collaborative filtering style recommendation system in which recommendations are generated based on crowd-sourced privacy expectations [62]. The authors test this system and find that it generates recommendations which match users' privacy expectations. In a separate effort, Liu B. et al. design and implement a mobile privacy assistant app for Android devices capable of generating recommendations for app permission settings [61]. As part of their study, the authors develop a preference model based on data from user study. In testing, the authors find that around 80% of resulting permission recommendations were accepted by participants during the study.

3.2.2 Explanations for Recommendation Systems

In parallel with the development of recommendation systems, research has sought to develop explanations for automated recommendations. Papadimitriou et al. propose a taxonomy of different 'styles' of explanations consisting of social ('human') style explanations, item style explanations, and feature style explanations [79]. The authors describe how these styles of explanations incorporate information about other users, choices made by users about similar items, and features previously rated by users respectively. In a later study, Svrcek et al. extend these concept by proposing a class of 'hybrid' style explanations which combine elements of two or more explanation styles into a single explanation [97]. The authors test these explanations in a user experiment, finding that 'hybrid' explanations improved users' attitudes about recommendation systems. The authors further suggest that by varying the selection of explanation styles, it may be possible to personalize explanations.

Several studies have examined the effectiveness of 'feature' style explanations. In an early paper, Billsus et al. develop content-based explanations for recommendations produced by an automated news agent [15]. These explanations link news article recommendations to previous user choices. In testing, the authors find that the explanations are useful as a tool for allowing users to provide feedback for the recommendations and improving their accuracy. Later, Cramer et al. test how varying the level of transparency within explanations impact the adoption of art recommendations [24]. These recommendations were generated based on stated user preference. The authors find that explanations which provided users insight into how the recommendation was made increased adoption rates.

Previous studies have also examined the effectiveness of social style explanations. In an early paper, Herlocker et al. examine social explanations for recommendations that are based on collaborative filtering [44]. Through a users study, the authors test different explanations designs for a movie recommendation system, finding that the explanations increased the acceptance of the automated recommendations. In a more recent paper, Sharma et al. test the effectiveness of social style explanations which incorporate information about other users' preferences for music recommendations [90]. While the authors find that the explanations prompt users to consider a recommended artist, there was a lesser impact on users' later artist ratings.

More recently, researchers have examined ways to evaluate explanations for automated recommendations. Tintarev et al. propose a framework for evaluating explanations based on the goals system designs may have for implementing explanations [99]. These goals include explaining how the recommendation system works (Transparency), helping users make better decisions (Effectiveness), and helping users arrive at decisions more quickly (Efficiency). In a related study, Tran et al. explore the needs expressed by users regarding explantions for a movie recommendation system [101]. The authors find that users want to see explanations more when they are not satisfied with the recommendation. They also find that many of the goals of explanations are interdependent and impact overall perceptions of satisfaction in the recommendation system.

To our knowledge, no study exists that examines explanations in the context of mobile permission recommendations. Given the complexities of privacy decision making and the emergence of recommendation systems for mobile privacy settings, addressing this gap in the literature may yield insights that could improve privacy recommendation systems and help users achieve better privacy outcomes.

3.3 Study Design

We study explanations in the context of recommendations for hypothetical privacy disclosure choices. Specifically, we designed a survey wherein we examine three 'styles' of explanations for a hypothetical recommendation. Within the survey, participants are asked whether or not they would adopt a recommendation from a hypothetical privacy assistant app. The recommendation is to either allow or deny a hypothetical video conferencing app called 'VideoMeet' access to their location depending on their responses to several privacy questions. We randomly assign participants to one of four experimental conditions which varies the amount and type of information presented about the decision. In all conditions, participants are presented with a recommendation. In three of the conditions participants are additionally presented with and explanation alongside the recommendation. These explanations vary in style and content. After reviewing the scenario, recommendation, and possible explanation, participants then decide whether or not they would adopt the recommendation in the context of the privacy decision. Following the privacy decision, participants answer a number of followup questions designed to measure their perceptions and preferences regarding different elements of the intervention. Details of the recommendation, experimental conditions, explanations, and followup questions are provided below.

3.3.1 Privacy Questions

To help generate the recommendation and provide content for one of the later explanation conditions, we begin the survey by asking participants a series of five ‘Privacy Questions’. These questions mirror the style of privacy questions asked within the mobile privacy assistant app for Android devices developed by Liu et al. [61]. Each question asks participants if they are comfortable with a particular category of app accessing a permission type. The app category and permission names are drawn from Android ecosystem. The selection of app category and permission pairs are the same for all participants. Participants respond to the privacy questions by indicating their comfort along a six-point Likert scale ranging from ‘very comfortable’ to ‘very uncomfortable’.[1] By using an even number of response options, we force participants to state a preference in in one direction or the other.

After participants finish responding to the five privacy questions, they are shown a screen with the text “Examining your responses to the privacy questions.’ Participants must remain on this screen for five seconds before progressing in the survey. This delay simulates the back-end processing that a real privacy assistant may perform and may increase the sense among participants that the hypothetical recommendation which follows is based in part on their responses.

3.3.2 Privacy Recommendation

Following the privacy questions, participants were presented with a recommendation from the hypothetical privacy assistant app. This recommendation was to either allow or deny access to a hypothetical video-conferencing app called ‘VideoMeet.’ When introducing the hypothetical ‘VideoMeet’ app, we inform participants that the app is

[1] See Appendix B.2 for the full text of the privacy questions.

requesting location access to notify them about nearby contacts and enable targeted advertising. The text of the recommendation is given below:

> (Allow Recommendation) "We recommend that you **ALLOW** the **VideoMeet** app access to your **Location**."
>
> (Deny Recommendation) "We recommend that you **DENY** the **VideoMeet** app access to your **Location**."

We assign participants to either an 'allow' or 'deny' recommendation based on their responses to the privacy questions. We first score their responses by counting the number of scenarios with which they report being generally comfortable. If a participant indicates they are generally comfortable with a majority of the scenarios in the privacy questions (3 or more), they are assigned and 'allow' recommendation. Likewise, if a participant is generally uncomfortable with a majority of the privacy question scenarios, they receive a 'deny' recommendation. Our goal in assigning recommendations is to increase the perceived quality of the recommendation, potentially reducing the likelihood that the recommendation would strongly conflict with existing privacy preferences. Because the decision to 'allow' location access differs from the decision to 'deny' in terms of privacy endowment (and because participants that self-select into the 'deny' recommendation may be more privacy sensitive), we treat the recommendations separately in our analysis.

3.3.3 Experimental Conditions

After they are presented with the recommendation, participants are randomly assigned to one of four experimental conditions. These conditions vary the information that is presented to participants within the privacy decision. Details on each of the four conditions are provided below.

1. No Explanation - Participants in this condition are only presented with the hypothetical privacy permission recommendation with no explanation.

2. Social Explanation - In addition to the recommendation, participants are shown an explanation that references the choices of other users.

3. Behavioral Explanation - Participants in this condition are shown the recommendation along with an explanation that references their response to one of the privacy questions.

4. Hybrid Explanation - Drawing from Svrcek et al., we develop a 'hybrid' explanation by combining the social and behavioral style explanations into a single explanation [97].

Drawing from the literature on explanations, we may expect the likelihood of participants adopting the hypothetical recommendation to be higher for the three explanation conditions relative to the 'no explanation' condition [100]. Likewise, we may expect the hybrid explanation to increase the likelihood of recommendation adoption relative to the social and behavioral explanations if we assume that the effects of the explanations are somewhat additive [97].

3.3.4 Explanations

While each of the three explanations we test in this study provided users with additional information about the recommendation, they reflect different styles of explanations as defined by Papadimitriou et al. and leveraged different information about the hypothetical recommendation system [79]. Specifically, we develop We modeled the hypothetical recommendation system on the mobile privacy assistant developed by Liu et al. which uses previously collected user data and asked similar privacy questions to generate recommendations [61]. Each of the social, behavioral, and hybrid explanations are described in more detail below.

The social explanation builds upon Sharma et al. providing participants with insight into how the recommendation was generated by making them aware that users with potentially similar preferences adopted the recommendation made by the hypothetical privacy assistant [90]. Participants may be more inclined to adopt the recommendation if they are aware of others with similar preferences which made the same decision [96]. The text of the social explanation is shown below.

> (Allow Recommendation) "Users whose answers to the Privacy Questions are similar to yours allow conferencing apps like VideoMeet access to their Location"

> (Deny Recommendation) "Users whose answers to the Privacy Questions are similar to yours deny conferencing apps like VideoMeet access to their Location"

In contrast, the behavioral explanation is modeled on 'feature' style explanations, and was intended to inform participants about particular decisions they made that influenced the recommendation [79, 24]. We may expect these explanations to increase recommendation adoption by providing further transparency into how the hypothetical recommendation system works. In the context of the hypothetical scenario, we informed participants that their answers to the Privacy Questions influenced the resulting recommendation. The text of the behavioral explanation is shown below.

> (Allow Recommendation) "Your answers to the Privacy Questions suggest that you may be comfortable allowing conferencing apps like VideoMeet access to your Location"

> (Deny Recommendation) “Your answers to the Privacy Questions suggest that you may not be comfortable allowing conferencing apps like VideoMeet access to your Location”

The hybrid explanation combined the information from the social and behavioral explanations into a single explanation. We presented the hybrid explanation to participants as a bulleted list containing the text of the social and behavioral explanations.

3.3.5 Quality and Preference Questions

The remainder of the survey is composed of questions designed to control for the quality of the recommendation and explanations, along with participants’ attitudes regarding the hypothetical mobile privacy assistant and online privacy generally. Together, these questions facilitate our primary regression analysis along with potential followup analyses. We describe these questions in detail below. The full text for all of the followup questions can be found in Appendix B.3.

Memory Recall: We included three memory recall questions immediately after the privacy decision which ask participants to recall details of the hypothetical scenario and experimental intervention. Specifically, we asked participants to recall the name of the hypothetical app (‘VideoMeet’), the permission requested by the app (‘Location’), and whether the recommendation suggested allowing or denying the recommendation. Each of these questions was multiple choice with a single correct answer. By counting the number of correct responses, we created a recall score for each participant. This score was later used as an exclusion criteria in our analysis [2]. The memory recall questions are intended to ensure that participants were more

[2] We additionally tested a fourth memory recall question where participants were asked to recall the purposes for which the ‘VideoMeet’ app was requesting access to their location. This question was

likely to have paid attention during our primary experimental intervention and considered the details of the intervention when making their choice about adopting the hypothetical recommendation.

Recommendation Quality: Following the memory recall questions, we asked participants a series of questions designed to measure their perceptions of the quality of the recommendation. While different metrics exist for assessing the quality of recommendations, 'accuracy' is among the most frequently used [39]. We define recommendation accuracy within the context of our study to be the extent to which adopting the recommendation would result in an outcome that would better align with participants' privacy preferences. We measure this concept with a single question along a Likert scale. We use participants' response to this question as the explanatory variable for recommendation quality within our regression models. Following this question, participants are asked additional questions about how the recommendation agreed or conflicted with their privacy preferences and whether they thought location access was necessary for the VideoMeet app. These followup questions take both multiple choice and open-ended formats. Prior to answering any of the recommendation quality questions, participants are shown the hypothetical recommendation from the study intervention again to refresh their memory.

Explanation Quality: We evaluate the quality of the social, behavioral, and hybrid explanations using questions designed to measure how well each explanation achieves three of the explanatory goals identified by Tintarev et al. [100]. In total,

multiple choice, but allowed for participants to select multiple responses corresponding to different purposes. In total, there were two correct selections out of six, corresponding to the two purposes for location access provided in the hypothetical scenario. When tested, few participants answered this question correctly by selecting *only* the two correct purposes. We therefore did not use the results from this question in compiling the recall score.

we asked three questions measured along Likert scales to capture how well the explanations increased transparency in how the hypothetical recommendation system worked, helped lead to satisfactory outcomes, and made the decision process more efficient. We take the average of these responses to construct the explanatory variable for explanation quality used in our regression models. We additionally measure the extent to which understanding how the privacy assistant generated the recommendation was important to the participant, the importance of the explanation to the adoption decision, and the factors that influenced their decision whether or not to adopt the recommendation. Prior to answering any of the explanation quality questions, participants are shown the explanation from the study intervention again to refresh their memory.

Trust in Privacy Assistant: Because trust in the recommendation system is identified in the literature as an important factor in users' decisions of whether or not to adopt a recommendation, we ask several questions to assess participants' perceived trust in the hypothetical privacy assistant app. Our questions are adapted from McKnight et al., and measure participants' perceptions of the hypothetical privacy assistant's honesty, benevolence, and competence [70, 69]. Each of these questions is measured along a Likert scale. We use the average of these responses for each participant as our trust explanatory variable.

Video Conferencing App Preferences: Existing preferences for real video conferencing apps participants may have installed on their mobile device (such as Zoom, Microsoft Teams, Webex, etc.) may also impact their decision of whether or not to adopt the hypothetical recommendation. Because of this, we capture which video conferencing apps participants have installed on their mobile device, the importance of these video conferencing apps to participants' day-to-day responsibilities, and whether they think the location access is necessary for the function of the video conferencing apps. We use whether participants have at least one video conferencing app installed on their mobile device as an explanatory variable.

Privacy Attitudes: As a final component of the survey, we measured participants' general attitudes regarding online privacy using the Internet Users' Information Privacy Concerns (IUIPC) scale [65]. This scale consists of ten privacy statements for which participants indicate their agreement along a Likert scale. These responses are averaged into a single score which we used as an explanatory variable in our regression analysis. In addition, we captured participants' general comfort regarding location-based targeted advertising along a Likert scale for potential followup analysis.

We pre-registered this study design along with our analysis plan on Open Science Framework prior to collecting data.[3]

3.4 Results

3.4.1 Demographics

We recruited a sample of $1,098$ participants from Prolific[4] to complete our survey in January 2022. Our sample size was informed by a post-hoc power analysis conducted using data from previous pilot studies. From this sample, we excluded 155 participants that answered one or more of the memory recall questions incorrectly. This allowed us to screen out participants that may not have paid attention during the study intervention. We additionally excluded 1 participant with incomplete demographic

[3] `https://osf.io/hrt7n`

[4] `https://www.prolific.co/`

data. We perform our analysis on a final sample of 942 participants.[5] Within our final sample, 48% (457) of participants were male and 52% (485) were female. The mean age of participants was 32 years with a median age of 28 years. 49% (465) of participants had completed a Bachelor's degree or higher.

3.4.2 Summary Statistics

As described in Section 3.3.2, we partition our sample based on the recommendation (either 'allow' or 'deny') that participants received. In total, 39% (372) of participants received an 'allow' recommendation while 61% (570) of participants received a 'deny' recommendation. Among those that received an 'allow' recommendation, 51% (188) responded that they would accept the recommendation while 49% (184) responded that they would reject the recommendation. Among participants who received a 'deny' recommendation, 95% (540) stated that they would accept the recommendation while 5% (30) stated that they would reject the recommendation. While the proportion of participants that would accept the 'allow' recommendation is roughly equal to those that would reject the recommendation, there is a large disparity in the acceptance rate for the 'deny' recommendation, with nearly all participants stating they would accept the recommendation.

Table 3.1 shows the rates and counts at which participants stated they would accept and reject the 'allow' and 'deny' recommendations. For the 'allow' recommendation, the behavioral explanation has the largest difference in adoption rate relative to the no explanation condition. However, this difference is not significant when measured using a Chi-squared test, with a p-value of $p = 0.31$. The 'deny' recommendation only shows minor variation across conditions, with the largest difference

[5] Our analysis for each study was conducted in R using publicly available packages for regression analysis and data presentation [35, 36, 46, 110].

Table 3.1: Adoption Intent by Condition

	Allow Recommendation			Deny Recommendation		
Condition	Accept	Reject	Rate	Accept	Reject	Rate
No Explanation	45	50	47%	137	4	97%
Social Explanation	46	50	48%	133	6	96%
Behavioral Explanation	54	43	56%	133	10	93%
Hybrid Explanation	43	41	51%	10	137	93%

occurring between the no explanation condition and the behavioral and hybrid explanation conditions. Neither of these differences are significant when measured with a Chi-squared test (with p-values of $p = 0.18$ and $p = 0.20$ respectively).

Table 3.2 and Table 3.3 display the summary statistics for our primary explanatory variables. The statistics for the recommendation, trust, and IUIPC scores are calculated using data from the full sample of participants while the explanation quality score is calculated from the sub-sample of participants randomly assigned to one of the three explanation conditions.

Table 3.2: Explanatory Variable Summary Statistics for Allow Recommendations

Variable	Mean	St. Dev.	Range
Rec: Quality Score	2.844	1.227	1 – 5
Exp: Quality Score (sub-sample)	3.099	1.177	1 – 5
Trust Score	3.387	1.024	1 – 5
IUIPC Score	5.805	0.870	1 – 7

Among participants that receive an allow recommendation, the mean recommendation quality score is below the mid-point of 3, suggesting that they tend to not agree that the recommendation would lead to an outcome better aligned with their privacy preferences. Conversely, the mean score for the composite explanation quality variable is near the mid-point (only slightly above it). This suggests that participants on average do not strongly agree nor disagree that the explanation achieves the three sub-goals from Tintarev et al [100]. Finally, the trust and IUIPC scores are both high, suggesting that participants tend to trust the capability of the hypothetical privacy assistant and that they tend to be privacy sensitive.

Table 3.3: Explanatory Variable Summary Statistics for Deny Recommendations

Variable	Mean	St. Dev.	Range
Rec: Quality Score	4.292	1.061	1 – 5
Exp: Quality Score (sub-sample)	3.479	1.052	1 – 5
Trust Score	4.032	0.824	1 – 5
IUIPC Score	6.350	0.690	1 – 7

Overall, the mean scores for each of the variables are higher for participants that receive 'deny' recommendations relative to those that receive 'allow' recommendations. This correlates with increased perceptions of quality, trust, and privacy sensitivity. In the case of recommendation quality, the mean score for participants that receive a 'deny' recommendation is now above the mid-point. This suggests that they tend to agree that following the recommendation would lead to an outcome more in line with their privacy preferences. While the the average IUIPC score is high across both tables, the higher value for participants that receive a 'deny' recommendation is likely the result of self-selection – where recommendations are assigned based on participants' responses to questions about their preferences regarding mobile privacy permissions. Indeed, the number of scenarios from the privacy questions that participants report being comfortable with exhibits a moderate inverse correlation to IUIPC scores with a Pearson's r of -0.42.

Finally, most participants had at least one video conferencing app (such as Zoom, Microsoft Teams, or Cisco Webex) installed on their mobile device – with 94% of users assigned to the 'allow' recommendation and 85% of users assigned to the 'deny' recommendation.

3.4.3 Logistic Regression

We constructed two logistic regression models to examine the effectiveness of explanations. In each of the models, the dependent variable of interest is participants' stated intention ('intent') for whether they would adopt the hypothetical mobile permission recommendation. The explanatory variables control for perceptions of recommendation quality, explanation quality, existing preferences for video conferencing apps, online privacy preferences, and demographics including education and sex. The two models differ only in terms of the subset of the sample employed and the research question(s) examined. Model 1 examines whether the presence of explanations for the recommendation significantly impacts the log-likelihood of participants intent to adopt the recommendation. The specification for this model is shown in Equation 3.1 below.

$$\begin{split}\frac{P(adoptIntent)}{1 - P(adoptIntent)} = \beta_0 + (\beta_1 * condition) + (\beta_2 * recQuality) + (\beta_4 * appPrefs) \\ + (\beta_5 * privPrefs) + (\beta_6 * demographics) + \epsilon\end{split} \tag{3.1}$$

Where Model 1 uses data from all participants in the sample, Model 2 considers only those participants assigned to one of the three explanations conditions. This allows us to compare the social, behavioral, and hybrid explanations relative to each other. The specification for Model 2 is shown in Equation 3.2 below.

$$\frac{P(adoptIntent)}{1 - P(adoptIntent)} = \beta_0 + (\beta_1 * condition) + (\beta_2 * recQuality) + (\beta_3 * expQuality) + (\beta_4 * appPrefs) + (\beta_5 * privPrefs) + (\beta_6 * demographics) + \epsilon \tag{3.2}$$

Table 3.4 shows the coefficient estimates with standard errors for Model 1.[6] None of the three explanation conditions exhibit significant effects at the $p < 0.5$ level. Only the effect of the hybrid explanation is weakly significant at the $p < 0.1$ level for the 'Deny' recommendation, however this result is not robust when participants with lower memory recall scores are included in the sample. Across both 'allow' and 'deny' recommendations, the coefficients for recommendation quality and privacy attitudes (IUIPC) are highly significant. The directions of these coefficients make intuitive sense, as they suggest that higher perceptions of recommendation quality lead to an increased log-likelihood of adopting the recommendation and that higher privacy sensitivity lowers the log-likelihood of adopting the 'allow' recommendation while raising the log-likelihood of adopting the 'deny' recommendation. Likewise, the coefficients for trust and the sex of participants were significant for the 'allow' recommendation. In the case of the former, the coefficient suggests that participants with higher trust in the capability of the hypothetical recommendation system are more likely to adopt the recommendation. Finally, age was significant for the 'allow' recommendation, but only at the $p < 0.1$ level.

Table 3.5 displays the coefficients along with standard errors for Model 2.[7] We do not find significant effects for the social and behavioral explanation conditions relative to the hybrid explanation. Many of the remaining results are similar to

[6] The baseline condition for this model is a hypothetical female participant with no formal education that is assigned to receive no explanation.

[7] The baseline condition for this model is a hypothetical female participant with no formal education that is assigned to receive the hybrid recommendation.

Table 3.4: Effects of Explanations

	Dependent variable:	
	intent	
	Allow	Deny
	(1)	(2)
Social Explanation	−0.337	−0.355
	(0.421)	(0.691)
Behavioral Explanation	0.294	−1.022
	(0.434)	(0.630)
Hybrid Explanation	0.185	−1.111*
	(0.429)	(0.647)
Rec Quality	1.226***	0.524***
	(0.157)	(0.153)
Trust	0.878***	−0.046
	(0.183)	(0.243)
Video Conf Apps Installed	0.129	−0.517
	(0.642)	(0.668)
IUIPC	−0.685***	0.464**
	(0.194)	(0.207)
Age	0.023*	0.010
	(0.013)	(0.018)
Male	0.668**	−0.386
	(0.326)	(0.420)
Doctorate Degree	−13.012	15.644
	(882.744)	(910.972)
Graduate Degree	−12.661	1.128
	(882.744)	(1.281)
High School Diploma	−12.300	1.594
	(882.743)	(1.225)
Secondary Education (GED)	−13.889	0.532
	(882.744)	(1.602)
Technical/Community College	−13.020	1.201
	(882.744)	(1.238)
Undergraduate Degree	−12.400	1.660
	(882.743)	(1.228)
Constant	8.836	−2.316
	(882.745)	(2.049)
Observations	372	570
Log Likelihood	−146.382	−103.532
Akaike Inf. Crit.	324.764	239.065
Note:		*p<0.1; **p<0.05; ***p<0.01

those in Table 3.4, where the coefficients for recommendation quality, trust, IUIPC score, and participant sex sharing similar direction and significance. In contrast to Model 1, the IUIPC coefficient is only significant at the $p < 0.1$ level for the 'deny' recommendation.

An addition to Model 2 is the coefficient for explanation quality. This is significant at the $p < 0.05$ level for the 'allow' recommendation, suggesting that higher perceptions of the quality of the explanation were associated with a higher log-likelihood of adopting the recommendation overall. This does not comment on the performance of the explanations relative to each other.

We used generalized additive models to detect omitted transformations of the independent variables. This analysis indicated the each of the independent variables was roughly linear. Additional model specifications with incremental control variables are provided in Appendix B.1.

3.4.4 Followup Analysis

We examined a subset of additional factors which may provide further context to the null effects we observe on our primary experimental manipulations. When asked to describe the factors most important to their decision whether or not to adopt the hypothetical recommendation, many participants noted that they took into consideration whether they thought the 'VideoMeet' app needed to access their location while others described concerns about privacy and targeted advertising. Permission necessity was the most common theme in the open-ended responses. One participants described their concerns by stating, "The biggest factor is that I can't see a reason why it needs my location for the meeting". Another participant stated, "I didn't think this was necessary information for the app to have". These sentiments, when

Table 3.5: Comparison of Explanations

	Dependent variable:	
	intent	
	Allow	Deny
	(1)	(2)
Social Explanation	−0.479	0.654
	(0.417)	(0.611)
Behavioral Explanation	0.141	0.028
	(0.426)	(0.510)
Rec Quality	1.125***	0.547***
	(0.176)	(0.168)
Exp Quality	0.425**	−0.108
	(0.192)	(0.258)
Trust	0.713***	−0.077
	(0.232)	(0.316)
Video Conf Apps Installed	0.221	−1.071
	(0.787)	(0.814)
IUIPC	−0.624***	0.435*
	(0.230)	(0.223)
Age	0.015	0.003
	(0.015)	(0.019)
Male	0.911**	−0.365
	(0.378)	(0.460)
Doctorate Degree	−13.598	16.149
	(882.744)	(964.921)
Graduate Degree	−13.583	1.845
	(882.744)	(1.396)
High School Diploma	−13.117	1.852
	(882.744)	(1.286)
Secondary Education (GED)	−13.861	0.342
	(882.744)	(1.707)
Technical/Community College	−13.680	1.603
	(882.744)	(1.324)
Undergraduate Degree	−13.347	2.517*
	(882.744)	(1.325)
Constant	9.043	−2.534
	(882.746)	(2.148)
Observations	277	429
Log Likelihood	−110.528	−84.711
Akaike Inf. Crit.	253.056	201.421
Note:		*p<0.1; **p<0.05; ***p<0.01

held by participants, may contribute to stronger preferences against the disclosure that lower the likelihood of participants being influenced by a recommendation – and subsequently accepting that recommendation if it disagrees with those preferences.

In addition to concerns about permission necessity, concerns about privacy were also frequently expressed by participants. One participant stated, “The targeted ads was the factor that lead me to choose no and nearby contacts isn’t necessary for the function of this app”. Another participant responded, “Targeted ads are annoying and based on location”. These privacy concerns may further decrease the likelihood of participants considering the recommendation if they increase the disparity between perceived benefits and drawbacks to disclosure.

When we later asked participants to quantify how necessary they thought it was for the hypothetical video conferencing app to access their location, those assigned to both the ‘allow’ and ‘deny’ recommendations tended to respond that it was only ‘slightly necessary’ or ‘not at all necessary’ (responses to this question were measured along a 5-point Likert scale ranging from ‘not at all necessary’ to ’extremely necessary’). On average, participants that received a ‘deny’ recommendation tended to view the permission as less necessary for the video conferencing app to have access to while in use than participants that received an ‘allow’ recommendation (with mean scores of 1.413 and 2.064 respectively). Overall, many participants appear to have weighed the details of the hypothetical scenario and determined that the ‘VideoMeet’ app did not need access to their location to provide needed functionality. Given that participants also exhibited high privacy sensitivity on average, the combination of low perceived ‘benefit’ to location disclosure paired with high privacy sensitivity may have created strong preferences among participants to either ‘allow’ or ‘deny’ access to the location permission. In turn, this may have lowered the impact of the recommendation to participants’ decisions.

We asked participants similar followup questions regarding explanation quality. When asked to describe how the explanation impacted their decision, few considered it important to whether or not they adopted the recommendation. One participant found the content of the social explanation not important to their decision stating, “I didn’t really care about what other people chose in this situation, since I felt like I had to make a different choice for my own privacy”. Participants also largely viewed the explanations as either confirming their decision-making (where they agreed with the recommendation) or ‘interesting’ but ultimately not a factor in their decision (where they disagreed with the recommendation). In describing their decision to accept a ‘deny’ recommendation, one participant stated, “Somewhat important, but I knew that i would not be comfortable with it anyway”. These themes are reflected in a closed-ended question where participants rated the importance of the explanation to their decision on a 5-point Likert scale ranging from ‘Not at all important’ to ’Extremely important’. Participants generally rated the explanations as not important, however participants that received a ‘deny’ recommendation rated the explanation as slightly more important than those that received an ‘allow’ recommendation (with mean scores of 2.695 and 2.451 respectively).

3.5 Discussion

Overall, we do not find evidence showing that the explanations tested impact the likelihood of adoption for the hypothetical mobile privacy recommendations. These results do not prove the lack of an effect, only that our study did not identify such effects for explanations if they are truly present. Further studies are warranted to either prove or disprove their existence. Instead, the pairing of the results from our regression analysis with our followup analysis may suggest a limit to the effects of explanations (if truly present) in the face of strong privacy preferences. Our followup analysis of the perceptions of the recommendations and explanations revealed that participants were largely concerned with the purpose for which the 'VideoMeet' app would access their location data and further found the ability of the 'VideoMeet' app to access their location to be unnecessary. Paired with an overall high privacy sensitivity, the disparity between perceived risks and benefits within the hypothetical scenario may have translated into a lower likelihood of participants considering the recommendation if that recommendation disagreed with their preferences regarding the disclosure. Our results therefore may suggest that explanations cannot improve the likelihood of adoption for recommendations if the recommendations are not of good quality.

To be successful, future studies which seek to examine the impact of explanations on recommendation adoption should take care to minimize the disparity between perceived risks and benefits within the disclosure scenario and ensure that the perceived quality of the recommendations is high. Within our study, limitations in the design of our disclosure scenario, recommendations, and explanations may have contributed to these perceptions among participants and therefore potentially reduced the likelihood of identifying effects for explanations if they were truly present. Many of these limitations stemmed from the hypothetical nature of our study, where we recorded participants' intentions rather than their real behavior. Previous studies have found that hypothetical decision problems may change how participants value risks and benefits, and activate different cognitive biases compared to real decision problems [54, 106]. It is possible that a study of explanations for real privacy recommendations would identify stronger and significant effects.

Additionally, participants in a hypothetical scenario may make different assumptions about the risks and benefits of disclosure that would not be present in a study on real behavior. While we defined these risks and benefits within our scenario description and captured preferences for existing video conferencing apps, it is still possible that participants made additional assumptions about the functionality of the 'VideoMeet' app based on their experiences with other video conferencing apps that we did not account for. A study of real behavior would avoid this limitation. While we initially designed and built the infrastructure for a field experiment to examine explanations in the context of real mobile privacy recommendations, we were forced to abandon the approach after the quality of the underlying recommendations was too low.

Within the context of our hypothetical study, we designed the disclosure scenario based on existing mobile privacy recommendation apps [61]. While we refined the design of our hypothetical scenario over multiple pilot studies to make it realistic, it is possible that a scenario with different benefits and drawbacks to disclosure would elicit a more balanced rate of participants accepting and rejecting the recommendation.

Limitations in the design of our recommendations may also have reduced the likelihood of identifying effects for explanation on recommendation adoption intent if they were truly present. Within our hypothetical study, we assigned participants to either an 'allow' or 'deny' recommendation based on their answers to the privacy questions early in the survey. This simplistic model for user preferences may have contributed to perceptions of low recommendation quality among participants. Future studies of hypothetical or real behavior may wish to reexamine this approach with a more robust recommendation model.

Finally, our study examines explanations from a 'nudging' perspective. Each of the explanations that we test introduce new information at the time of the decision without forbidding any options or changing incentives – meeting the definition of a nudge as proposed by Thaler and Sunstein [98]. While this provides one frame for examining the effectiveness of explanations, they may also be evaluated from the perspective of helping users make informed privacy choices. Indeed, an effective explanation may be one which helps users recognize that the recommendation model used to generate the recommendation is not effective. While we measure both perceived trust in our hypothetical recommendation app and explanation quality for use as controls in our model, future studies may seek to examine outcomes other than adoption intent.

3.6 Conclusion

We conducted a study to examine whether three styles of explanations impacted the stated adoption intentions for hypothetical mobile permission settings recommendations. Participants completed an online survey in which they evaluated a hypothetical scenario containing a mobile privacy recommendation and decided whether or not they would adopt the recommendation. We variably presented participants with different explanatory information (including no explanatory information) for the recommendation across four experimental conditions. We constructed two logistic regression models to test whether the likelihood of participants' choosing to adopt the recommendation changed depending on the condition to which they were assigned. In Model 1, we examined whether the three explanation styles impacted the likelihood of recommendation adoption relative to no explanation. In Model 2, we assessed the performance of the explanations relative to each other. Across both models, we find robust effects for perceived recommendation quality and privacy preferences, but failed to identify effects for our experimental manipulations. A followup analysis suggests that strong preferences and uncertainty regarding the purpose of data collection may have contributed to our null effects. Many of these limitations are inherent to studies on hypothetical behaviors and future studies on explanations for automated privacy recommendations should consider focusing on real behaviors.

Chapter 4

Trends in Privacy Notices after the GDPR

Abstract

While multiple studies have examined changes in privacy notice designs on websites following the GDPR, less attention has been devoted to capturing the evolution of these changes over time. This study intends to address this gap in the literature by investigating the dynamics of privacy notice design between 2018 and 2020. To do this, we leveraged longitudinal data from a panel of 911 websites in the US and EU and examine the features of the privacy notices that appear on each website's homepage. We manually classified screenshots from each website using a rubric we developed of common privacy notice design features and assigned each screenshot one of seven labels indicative of the response to the GDPR. We use this data to examine whether website-level characteristics such as user location and traffic drive the adoption of particular response types over time and changes in the features from our rubric. We additionally examine the potential impact of four industry and government actions on response adoption and consent notice design. Our results identify geographic disparities in the prevalence of privacy notices in the US compared to the EU. We additionally identify a shift away from more coercive responses on US websites and

a shift towards consent responses on EU websites. We discuss the implications of these results, the limitations of our methodology, and possible directions for future research.

4.1 Introduction

Over the previous two chapters, I examined how two different types of behavioral interventions performed in experimental settings. To gain a broader understanding of how behavioral interventions may be presented to users in real settings and how that may be impacted by public policy, this chapter examines observational data collected from websites.

In 2016, the European Union enacted the General Data Protection Regulation (GDPR) with the goal of protecting the personal data of EU residents (GDPR Recital 1). To achieve this protection, the law introduced transparency and control requirements to allow data subjects to manage how personal data about them is stored and used. Many of these requirements are of interest to the privacy community due to the impacts they may have on consumer privacy. These impacts arise from changes made by the GDPR to how data is collected and processed. Under the law, data processors must now justify data collection under one of six legal bases at the time of collection (GDPR Article 6(1)). Among these allowed justifications is user consent, which aims to give users a greater level of control over how their data is used. Unique to these requirements is their extraterritorial scope, where the protections of the law apply to EU data subjects independent of their location (GDPR Article 3). Because of this, the new requirements had an immediate impact on any entity that collected or processed the data of EU data subjects.

With the introduction of the GDPR, many industries were forced to re-evaluate their data handling practices to meet the increased requirements for data protection and transparency [47, 21, 58]. In the case of websites, those which collected personal data for uses including user analytics and advertising were confronted with the task of providing a legal basis for the practice. How websites responded to this compliance burden may have been influenced by how website operators perceived the risk of potential enforcement actions such as fines resulting from different response types. This may have further varied by industry, with websites that relied on processing personal data of EU data subjects for a greater share of their revenue (such as in the case of content publishers that utilize behavioral advertising) responding differently compared to those that did not.

Leading up to and after the enforcement of the GDPR, many websites which collected data on users updated or introduced privacy notices to gain user consent for data processing [26]. However, adoption of consent as a response was not universal. Some websites sought other clauses under Article 6 of the GDPR such as 'legitimate interest' to justify their data collection practices (GDPR Article 6(1)(f)). Others still made minimal or no changes, choosing instead to keep older privacy notices or display none at all.

This fracturing of notice implementations among websites that process personal data may have been compounded by an initial lack of guidance by governmental bodies. While the GDPR lays out several requirements for consent, it does not provide any guidance for how it should be implemented. Although regulatory bodies such as the European Data Protection Board (EDPB) issued guidelines for compliance early-on, these documents were not legally binding and were often context-specific.

In the four years since its implementation, a strong consensus on what it means for websites to be GDPR compliant may only be starting to emerge [68]. Although justifications for data collection such as 'legitimate interest' have potentially fallen out of favor, consent notices that offer users differing levels of control over their data have appeared [32]. While some notices make it easy for users to reject data processing, others obstruct users by employing malicious designs or 'dark patterns' that implement nudges to make these choices difficult to exercise [76].

We leverage longitudinal data from a large panel of 911 'News and Media' websites from the US and EU to examine three research questions currently explored in the literature:

RQ1. What characteristics of websites (IP, Share EU Users, Alexa Reach, Wave Number) drive the adoption of different responses to the GDPR for US and EU websites?

RQ2. Do significant actions by DPAs or industry groups have an impact on websites' choice of response to the GDPR?

RQ3. Do the designs of consent notices (as defined by available actions) vary based on website characteristics (IP, Share EU Users, Alexa Reach, Wave Number) for US and EU websites?

We develop a rubric to classify the privacy notice features of $14,832$ screenshots observed over an 18 month period following the enforcement of the GDPR in May 2018. In addition to identifying the actions and messaging within privacy notices, we label each screenshot with one of seven 'response' labels which represent the way in which the website chooses to respond to the GDPR. For screenshots that contain consent notices, we further classify them based on whether they contain designs that nudge users to accept tracking. We use multinomial logistic regression models to examine the changes in the adoption of different response types and nudges within the screenshots in our sample based on website-level characteristics including visitor location, traffic, and time. We additionally use logistic regression to examine how the design of consent notices potentially change over time with respect to the same website-level characteristics. For both models, we additionally examine the potential impacts of four actions by government and industry that may have impacted the response of websites to the GDPR.

Overall, we observe a differential effect based on geography for the likelihood that a screenshot will contain privacy messaging. For both US and EU websites, visitors from the US are less likely to receive a privacy notice relative to US users. US websites are furthermore less likely to display privacy notices to users generally if they only receive a small share of their traffic from the EU. While we observe additional effects based on traffic and time, they do not consistently suggest websites are adopting a particular style of response based on these characteristics. However, we do observe a potential effect of two government and industry actions on the likelihood of 'strong nudges' appearing within consent notices for US and EU websites. Future studies may seek to expand upon this work by gathering more granular data on events, incorporating data on Consent Management Platforms, and capturing data which would facilitate an examination of consent notices that require multiple steps across several notice screens to reject tracking.

4.2 Related Literature

This study contributes to the growing literature on the GDPR and website privacy notice design by examining changes in privacy notices following the enforcement of the regulation.

Several studies have examined how websites have introduced or adapted privacy notices in response to the new requirements. Degeling et al. examined privacy policies and cookie notices for popular EU websites immediately before and after the GDPR [26]. While most websites had privacy policies before the GDPR, the authors find that many websites introduced privacy notices after the GDPR. In a related study, Utz et al. examined numerous privacy notices after the GDPR and tested how varying common UI elements impacted consent rates within a user study. [104]. The authors found that small changes in the display of notices impacted the rate at which users consented to tracking.

Additional studies have examined how changes to the UI elements of privacy notices can impact the rates at which users consent to tracking. Nouwens et al. examined implementations of the top five Consent Management Platforms (CMPs) across 10,000 websites [76]. Among their sample, the authors identified frequent uses of dark patterns and privacy notices with implied consent for tracking. In a small user study, the authors find that UI changes such as removing opt-out buttons have strong effects on consent rates. A similar study by Machuletz et al. tested how defaults along with the number of options available to users impact consent rates [64]. While the authors found the use of default accept buttons impacted the number of items users consented to, presenting users with more choices did not.

One previous study examined how these design choices can impact usability. Habib et al. conducted a lab study to examine users' ability to navigate common types of privacy notices and complete tasks such as opting-out of targeted advertising [41]. The authors find that while options to opt-out may be available to users, they are difficult to exercise in practice.

Studies have examined how the GDPR has impacted users' ability to manage their privacy. In a study of the privacy policies of FinTech firms based in Germany, Dorfelitner et al. find that the readability of privacy policies decrease after the GDPR while their length and quantity of data processed increases [27]. The authors argue that these changes may make true informed consent more difficult for users, in contradiction with the stated intentions of the GDPR.

Many of the existing studies on privacy notices within the literature form a picture about how different design choices impact users' consent choices. Common to many of these studies is the examination of privacy notices post-GDPR at one or two moments in time. Additionally, many of these studies only focus on privacy notices on websites when accessed from within the EU. Of those that focus on regional differences between privacy notices, Eijk et al. do not find that a majority of websites implement the same privacy notice for all users independent of location [31]. Separately, Sanchez-Rola et al. found that many websites continued to place tracking cookies on users' browsers by default shortly after the GDPR [86]. They additionally perform case studies on the privacy notice designs of three popular websites, identifying different mechanisms for users to reject tracking along with evidence of 'dark patterns'.

Only one study to date has examined changes in a subset of privacy notices over time [45]. In that study, Hils et al. examine implementations of consent notices from CMPs over time across a large panel of websites observed from both the US and EU. They focus in particular on the adoption of CMPs by websites, finding that adoption is most prominent among mid-ranking websites. They extend this analysis by breaking apart how the market share of CMP has shifted since the introduction of the GDPR, examining adoption relative to major GDPR enforcement events, and testing variations of CMP consent notice designs on real users. For the later item, they find that dialogs that require multiple steps for users to reject tracking take significantly longer than granting consent compared to notice designs that contain single-step options to reject tracking.

While in this chapter we perform similar analysis, we differ from Hils et al. in three significant ways. First, Hils et al. performs their main analysis on a sample of websites drawn from social media activity. This sample may be biased towards highly ranked or popular websites, potentially missing trends from lower ranked websites. In comparison, our sample (see Section 4.3) includes both high and low ranking websites from a variety of locations. Second, the primary focus of Hils et al. is CMP adoption among websites. It does not focus on consent notice design outside of a short users study. Our study focuses on the design features of the notices using more granular feature data, enabling us address broader questions of user behavior and intervention design. Finally, we leverage additional data on website traffic to examine whether these metrics predict websites' response to the GDPR.

While many existing studies demonstrate the shortcomings of privacy notices as they were when the studies were conducted, it is difficult to reflect on the regulation compliance more broadly without taking a more expansive view. Gaining a broader perspective on the adoption and implementation of privacy notices by tracking them over a longer period of time and from multiple locations may allow for a more nuanced assessment of the GDPR compliance among websites and may provide an additional input for policymakers considering similar privacy regulation in the US and elsewhere.

4.3 Study Design

To address our research questions, we leverage data from a parallel effort to examine the impact of the GDPR on content publishers [57]. Over the past four years, we have built a longitudinal data set which tracks a set of technical variables and content measures across a large panel of websites. This panel includes websites from the US and six EU countries: the UK, France, Germany, Italy, Spain, and the Netherlands [1]. For each of the countries in our sample, we gathered the 500 most popular websites

[1] We consider the UK to be part of the EU for this study as they implemented and enforced the GDPR.

as measured by Alexa Internet in early-2018. [2]. We supplemented this panel with a random sample of websites from Alexa's list of the top 1 million websites globally. This yielded a sample of 11, 254 websites from which we eliminated websites with few US or EU visitors, websites that provided adult content, and websites that were not assigned to any category. Our complete panel consists of 5, 474 websites.

At regular intervals or 'waves,' we visited each website in our sample using browsers instrumented with the OpenWPM framework from IP addresses in the US and EU. Through OpenWPM, we were able to collect technical data related to tracking for each website including cookies, HTTP requests, and the full HTML of the website. In addition, we captured screenshots of the homepages of each website for every 'wave' of data collection. Each screenshot contains a visual record of the website at the time of collection and includes any privacy notice a website may show users upon first loading the page. Because we visited each site from US and EU IP addresses, and visited each site 16 times between 2018 and the end of 2019, we have up to 32 screenshots for each website. We use these screenshots as the primary source of data for our analyses.

4.3.1 Screenshot Analysis

Within this sample, we visually examined the screenshots from the 1, 024 websites categorized by SimilarWeb as 'News and Media'. Unlike websites in other categories, 'News and Media' websites are more likely to rely on targeted advertising as a source of revenue [20, 37, 56]. Due to the reliance of targeted advertising on personal user data, these websites may therefore be more likely to display privacy notices to users regarding tracking and may be more likely to request consent.

For each screenshot in this sub-sample, we looked for the presence of privacy mes-

[2] https://web.archive.org/web/20180201030024/https://www.alexa.com/

saging in notices that may be displayed to users. Based on this messaging, we assign each screenshot to one of seven possible ‘responses’ that reflect the range of actions websites may take in response to the GDPR. These responses are: ‘No Notice’, ‘Block EU Users’, ‘Cookie Wall’, ‘Cookie Banner’, ‘Full Consent’, ‘Weak Nudge Consent’, and ‘Strong Nudge Consent’. We describe the definitions and methodologies for assigning each of these labels below.

4.3.2 Response Classification

When classifying a screenshot, we first determined if any privacy messaging was present. If the screenshot did not contain any privacy notice, we assigned it the response ‘No Notice’. This response indicates that the website either did not respond to the GDPR or was potentially justifying data collection through ‘legitimate interest’.

Of the screenshots that contained privacy messaging, we distinguished between those from websites which blocked EU users or forced consent through cookie walls from those that present other types of privacy notices. We classified a screenshot as ‘Block EU Users’ if it contained a notice to EU visitors that the website was unavailable - and no other content. Likewise, we labeled a screenshot as ’Cookie Wall’ if it forced users to consent to tracking before allowing them to view content. These screenshots would often contain a prompt to accept tracking in the center of the page with no option to reject tracking or dismiss the notice.

4.3.3 Privacy Notice Rubric

The remainder of privacy messaging within screenshots was contained within privacy notices such as pop-ups or banners. To distinguish between different types of privacy notices, we catalogued the features of the notice according to a rubric we developed by manually inspecting screenshots to identify common types of features and drawing from relevant academic literature on GDPR notice design [76, 87]. This rubric allows us to differentiate between different types of privacy notices by capturing the features of each notice along three dimensions: presentation, actions, and content.

The presentation feature concerns whether or not the notice obstructs the website in such a way that a user must address the notice before viewing the content of the website. We classify the presentation of each notice as either 'obstructing' or 'non-obstructing' using a single binary variable. The action features represent UI elements corresponding to different actions available to a user when interacting with the privacy notice. These include the presence of a button to 'accept' or 'allow' tracking, a button to decline tracking, an option for more information on cookie tracking, a settings button, a button to close the notice, a link to the website's privacy policy, and preference checkboxes to indicate the types of cookies that a user wishes to allow. We indicate the presence of each action in a notice using binary indicator variables. Finally, the content features concern the privacy messaging contained within the notice. These include whether the notice references tracking cookies or other types of cookies, whether the notice contains explicit options to opt-out of tracking, whether it only implies the ability to opt-out of tracking, whether it states that users will be tracked by default, and whether any preference checkboxes are pre-selected. As with the presentation and action features, the presence of each type of messaging is indicated with binary indicator variables.

Using the information from the rubric, we separate 'Cookie Banners' from notices which solicit consent from users. We define a cookie banner to be a privacy notice which informs users of the use of cookies, but does not offer them a means to opt-out. Unlike cookie walls, cookie banners do not explicitly force users to consent to tracking before allowing them to view the content of the website. Cookie banners often notified users that their consent for tracking was implied by their continued use of the website.

In contrast, we define a privacy notice to be a consent notice if it provides the user an option to reject tracking. Because variations in notice design can make the task of rejecting tracking more easy or more difficult for users, we split the response category of consent notices into three sub-categories to track the potential emergence of 'dark patterns' – nudges which may encourage users to consent to tracking by making it more difficult for them to reject tracking. Where a consent notice offers a button to clearly reject or decline tracking, we assign the response label 'Full Consent.' These types of notices are among the easiest for users to navigate and may be more in-line with the spirit of the GDPR [42]. Where a consent notice does not contain an explicit 'decline' button, but implies the ability to decline tracking and contains a 'settings' button through which users may access a secondary screen with cookie settings, we assign the response label 'Weak Nudge Consent.' Where no 'settings' button is present, but the ability to reject tracking is implied by the notice, we assign the response label 'Strong Nudge Consent'. We use the terms 'weak' and 'strong' in this context to refer the relative difficulty users may experience when attempting to reject tracking.

4.3.4 Traffic Variables

Because websites of different sizes may respond differently to the GDPR, we collect traffic measurements for each website at the time of each observation. Specifically, we use the 'Reach per Million' metric from the Alexa Internet ranking service [63, 91, 104, 94]. This metric can be interpreted as estimating the number of users per 1 million internet users that visit a website. Larger 'Reach per Million' scores translate into higher traffic for a website.

While the GDPR requires all websites that process the personal information of EU data subjects comply, websites may choose not to respond or respond in a way less likely to impact tracking and targeting practices if they receive little or no traffic from the EU. To account for this, we additionally collect the share of traffic each website in our sample receives from the EU according to SimilarWeb, a web-analytics service [3]. We refer to this metric as 'Share EU Users' within our data and record its value on a scale from 0 to 1. We collected this data once for each website at the beginning of our data collection period in early-2018. We assume that this value does not change significantly over the time period of the study.

4.3.5 Enforcement Actions and Guidance

Because of the panel structure of our data, we are uniquely able to assess the potential impacts of specific actions by government and industry on websites' responses to the GDPR as visible within screenshots. Specifically, we examine the potential impact of fines, guidance, and standards across four events that occurred during our window of data collection. In total, we selected four events based on their potential to have a significant impact on websites' responses. We list these events along with their corresponding dates in Table 4.1.

Table 4.1: Enforcement Actions and Guidance

Event	Date
Google CNIL Fine	Jan-21-2019
Dutch DPA Cookie Wall Guidance	Mar-7-2019
CNIL Cookie Guidance	Jul-23-2019
IAB TCF v2.0	Aug-21-2019

[3] `https://www.similarweb.com/`

Of these events, three are actions by data protection authorities while one is an action by an industry group. The earliest of these is the fine by the French Data Protection Authority (CNIL) against Google for failing to provide users with valid consent for ads personalization. Specifically, the CNIL observed that the information provided to users at the time of consent was not specific enough and that consent was not 'unambiguous' [16]. The resulting €50 million was one of the largest at the time and may have impacted the design of consent notices by making specific actions such as 'accept' and 'decline' buttons more available to users.

The second event we examine is the introduction of new guidance by the Dutch Data Protection Authority (DPA) regarding cookie walls. This guidance stated that the use of cookie walls was not compliant with the GDPR as they did allow a means for users to reject tracking [83]. We may expect to see the prevalence of cookie walls as a response within our sample of screenshots decrease following this guidance.

Similar to the Dutch DPA's cookie wall guidance, we additionally examine the potential impact of the CNIL's guidance on cookies. This guidance updated the CNIL's previous cookie guidance from 2013 and specified that websites cannot infer the consent of users for users scrolling or otherwise continuing to browse a webpage [59]. We might expect this guidance to potentially reduce the number of 'cookie banner' responses where the consent of users is implied as a result.

Finally, we examine the potential impact of the release of version 2.0 of IAB Europe's Transparency and Consent Framework (TCF) [34]. Notably, this update of the TCF incorporated feedback from publishers and offered them greater control over the advertisers they worked with and the legal bases they claimed for data processing [33]. If this standard was widely adopted following its introduction, we may expect to see an increase in consent notices and a potential decrease in older style cookie banners.

In the timeline of our data collection, both the CNIL Google fine and the Dutch DPA cookie guidance occurred between waves 10 and 11 in our data. Likewise, the CNIL cookie guidance and the introduction of the IAB TCF v2.0 occur between waves 13 and 14 in our data. Because of this, we code for the occurrence of these events using two binary indicator variables. We label screenshot observations that occur after the first two events (waves 11 and greater) as 'Events 1' and the screenshots observed after the second two events (waves 14 and greater) as 'Events 2'. Although we will not be able to separate the effects of the individual events in our analysis, we can potentially capture the effects of these enforcement and guidance actions more broadly.

4.4 Results

Our complete sample of classified screenshots contains $29,816$ observations [4]. From this sample, we exclude observations from websites that do not originate in the US or EU. We also exclude screenshots captured before the GDPR was enforced in May 2018 and observations where traffic data was missing. Following these exclusions, our final sample contains $14,832$ screenshots from 9 waves of data collection spanning the time period from late-May 2018 to November 2019.

We conduct our analysis on a final sample that covers 911 unique domains.[5] To account for high observed correlation between website location (US or EU) and a website's share of users originating from the EU, and to avoid potential issues resulting

[4] We performed data cleaning as an intermediary step where we screened for duplicate observations of the same screenshot, corrected errors resulting from the attribution of screenshots' domains in the database, removed screenshots collected with an ad-blocker enabled, and removed waves 11 and 15 where errors with OpenWPM data collection resulted in incomplete observations of the websites in our sample. We additionally matched the screenshot data with our data from OpenWPM, Alexa, and SimilarWeb. The processes combined reduced the number of screenshots from $40,854$ to $29,816$.

[5] Our analysis was conducted in R using publicly available packages for regression analysis and data presentation.[36, 46, 25].

from multicollinearity, we examine US and EU websites separately.

Out of the 911 websites in our sample, 444 are US websites and 467 are EU websites. Of the screenshots from these websites, $7,236$ originate from US websites while $7,596$ originate from EU websites. To determine the potential impacts of website-level characteristics and traffic metrics on the adoption and implementation of different types of consent notices, we additionally examine the subset of screenshots containing consent notices in isolation. This sub-sample consists of $3,586$ screenshots, with 749 screenshots from US websites and $2,837$ screenshots from EU websites.

4.4.1 Summary Statistics

Table 4.2 and Table 4.3 display summary statistics for each of the response types across two time periods at the beginning and end of our data collection window. Here, we define 'Early-GDPR' to be the first three waves of data collection after the GDPR was enforced (around May 2018) and 'Late-GDPR' to be the last three waves of data in our sample (up to November 2019). 'Screenshot Obs' represents the number of screenshots in our full sample that adopt each of the responses during the respective early and late time windows. 'Avg Share EU' reports the average share of EU users (between 0 and 1) among the screenshot observations for each response type and time period. Likewise, 'Avg Reach PM' reports the average 'Reach per Million' traffic measure among the screenshot observations for each response type and time period.

Across both tables, the proportion of 'No Notice' responses relative to other response types appears higher for US websites compared to EU websites. Furthermore, the proportion of 'No Notice' responses appears higher for US visitors relative to EU visitors. While we cannot make inferences about trends in response types from these tables alone, they provide a basis for understanding the regression models specified in Section 4.4.2 and Section 4.4.3.

Of additional note, we do not observe any screenshots with ‘Cookie Wall’ or ‘Full Consent’ responses for US websites when viewed from a US IP. Likewise, we only observe a small number of ‘Strong Nudge Consent’ and ‘Weak Nudge Consent’ screenshots for US websites when viewed from a US IP.

Table 4.2: Screenshot Response Characteristics: US IP

			Response						
			No Notice	Block EU Users	Cookie Wall	Cookie Banner	Full Consent	Weak Nudge Con.	Strong Nudge Con.
US Websites	Early GDPR	Screenshot Obs	1075	0	0	110	0	10	12
		Avg Share EU	0.03	-	-	0.05	-	0.32	0.16
		Avg Reach PM	200.29	-	-	217.52	-	198.77	489.77
	Late GDPR	Screenshot Obs	1085	0	0	90	0	19	13
		Avg Share EU	0.04	-	-	0.05	-	0.18	0.08
		Avg Reach PM	153.40	-	-	142.26	-	87.66	41.13
EU Websites	Early GDPR	Screenshot Obs	341	0	88	454	12	185	164
		Avg Share EU	0.75	-	0.96	0.81	0.70	0.83	0.85
		Avg Reach PM	244.83	-	36.45	232.04	39.60	212.45	128.19
	Late GDPR	Screenshot Obs	438	0	63	283	16	271	212
		Avg Share EU	0.75	-	0.96	0.81	0.84	0.77	0.86
		Avg Reach PM	157.74	-	21.00	122.59	247.63	216.10	70.34

Figure 4.1 and Figure 4.2 show how response choices change at the website level over the period of our study.[6] When observed from both US and EU IPs, the number of websites which adopt a ‘No Notice’ response appear to either hold steady or decrease over time. These effects generally appear to be more pronounced for EU websites when observed from an EU IP address. Additionally, websites in both figures which employ cookie banners seem to decrease while the number of websites with

[6] We present Figures 4.1 and 4.2 using interpolated values where our data contains missing observations. Where a website is missing an observation for a particular wave, we use the response classification of the last available wave of data. In the case where a website is missing an observation for wave 1, we use the response classification of the next wave for which there is an observation.

Table 4.3: Screenshot Response Characteristics: EU IP

			Response						
			No Notice	Block EU Users	Cookie Wall	Cookie Banner	Full Consent	Weak Nudge Con.	Strong Nudge Con.
US Websites	Early GDPR	Screenshot Obs	675	166	20	138	41	72	75
		Avg Share EU	0.03	0.01	0.05	0.05	0.05	0.10	0.06
		Avg Reach PM	135.58	55.74	1164.93	251.11	234.13	685.38	738.40
	Late GDPR	Screenshot Obs	619	177	13	130	64	92	107
		Avg Share EU	0.03	0.01	0.05	0.04	0.06	0.08	0.07
		Avg Reach PM	73.60	34.07	1236.33	103.89	151.51	512.73	368.19
EU Websites	Early GDPR	Screenshot Obs	267	0	85	488	27	192	186
		Avg Share EU	0.79	-	0.95	0.79	0.53	0.77	0.86
		Avg Reach PM	172.09	-	41.11	270.24	126.37	231.73	152.41
	Late GDPR	Screenshot Obs	222	0	64	292	31	381	290
		Avg Share EU	0.83	-	0.96	0.80	0.75	0.75	0.81
		Avg Reach PM	101.61	-	17.54	130.95	354.62	204.42	104.93

consent notices appear to increase. Overall, the trends appear to be stronger for EU observations relative to US observations. This may suggest that some websites may be deferentially applying privacy notices based on the location of the visitor. While these figures show how response types change in aggregate over time, we cannot tell whether or how individual websites may have switched between response types. We account for these dynamics within our regression models.

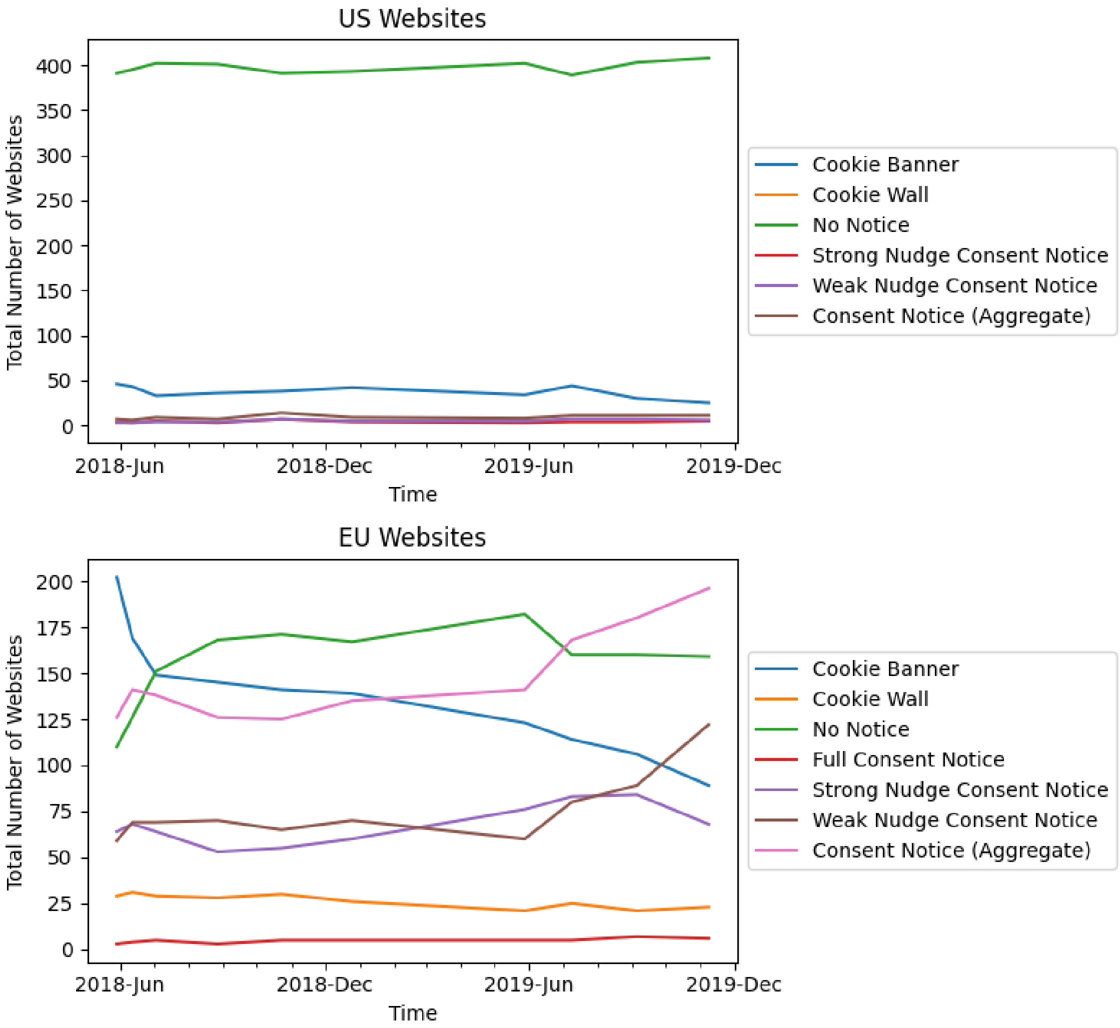


Figure 4.1: Counts of Website Response by Wave for US IP

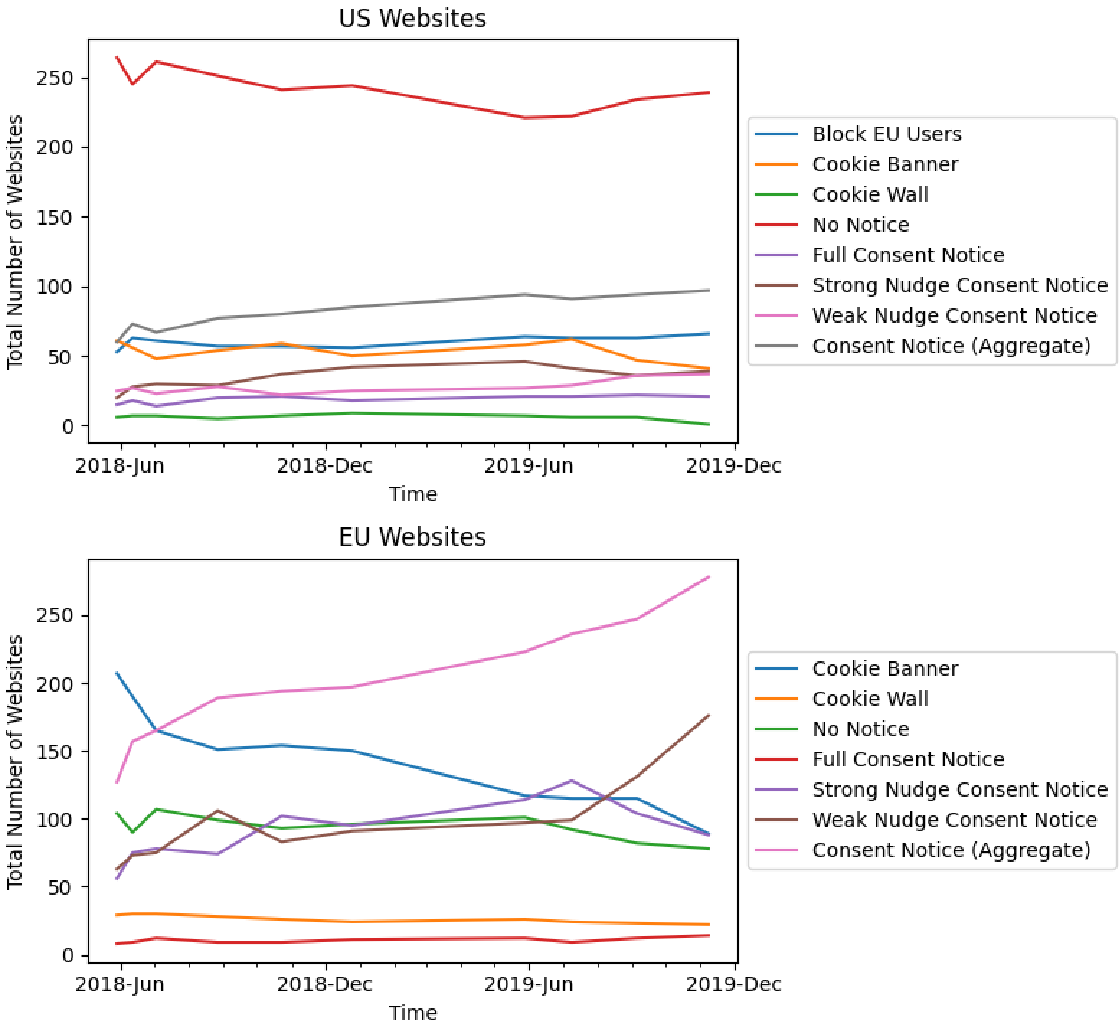


Figure 4.2: Counts of Website Response by Wave for EU IP

4.4.2 Multinomial Logitistic Regression

We use multinomial logistic regression to examine how website-level characteristics, traffic variables, and external events potentially impact the likelihood of screenshots adopting different response types. We construct two sets of models to examine both the full sample of screenshots and the sub-sample containing only consent notices. By examining the subset of screenshots with consent notices in isolation, we can examine how our control variables potentially impact nudge adoption. In both sets of models, we use the assigned ‘response’ as our dependent variable. Our control variables are the IP address of observation, share of EU users, reach per million, and our two binary ‘Events’ indicators. We additionally control for time fixed-effects including the wave of observation in our models.

Table 4.4 and Table 4.5 show the coefficient estimates with robust standard errors clustered at the website-level for US and EU websites respectively. [7]. The coefficient estimates in both models report the log-likelihood that a screenshot would adopt the given response relative to ‘No Notice’.

Across both tables and all response types, the coefficient estimates for US IP are negative and highly significant. This suggests that US visitors to US and EU websites may be less likely to be shown any type of privacy messaging regarding cookies or tracking - a finding which matches the descriptive observations in Figures 4.1 and 4.2. ‘Share EU Users’ is also highly predictive of response type for US websites, where all coefficients are highly significant. For all response types with the exception of ‘Block EU Users’, the sign of the coefficients are positive, suggesting that screenshots from US websites with higher shares of EU users are more likely to adopt these response types. As with the ‘US IP’ coefficient, this makes intuitive sense. US websites with higher shares of EU users may be more inclined to respond to the GDPR with some form of privacy messaging. Likewise, US websites with smaller shares of EU users (and therefore smaller shares of revenue from EU users) may be more likely to decide to block EU users rather than commit resources to implementing a GDPR response.

In contrast to US websites, only the ‘Cookie Wall’ and ‘Strong Nudge Consent’ response types are positive and significant for ‘Share EU Users’ for EU websites.

[7] Table 4.5 does not estimate effects for ‘Block EU Users’ responses as no EU websites blocked EU users.

Between these two, ‘Share EU Users’ has a much larger magnitude for ‘Cookie Wall’ responses. For EU websites with higher shares of EU users, adopting cookie walls may make sense for websites where the large majority of revenue from targeted advertising comes from EU users.

While we observe significant coefficient estimates for the ‘Events 1’ and ‘Events 2’ variables for cookie banners on US websites, these coefficients are of opposite signs and roughly equal magnitude. The cumulative effect may be that while there may be a small increase in the likelihood of cookie banner adoption after wave 11 (Events 1), this small increase dissipates after wave 14 (Events 2). It is difficult, therefore, to attribute this to any actions by EU governmental bodies as we would expect these effects to sustain over time. In contrast, both the ‘Cookie Wall’ and ‘Strong Nudge Consent’ response types have significant negative coefficient estimates for the ‘Events 2’ variable for US websites. The direction of these effects are in-line with what we might expect following the release of the CNIL cookie guidance and the IAB TCF v2.0 – which encourage explicit consent and streamline implementation respectively.

EU websites show positive significant coefficient estimates for ‘Events 1’ on ‘Cookie Banner’ and ‘Strong Nudge Consent’. ‘Events 2’ shows similar positive significant coefficient estimates for ‘Full Consent’ and ‘Weak Nudge Consent’ respectively. These potentially suggest an effect of government and industry actions on moving websites towards easier and more clear privacy messaging for users. That the likelihood of weak nudge and full consent notices increase following ‘Events 2’ relative to adopting ‘No Notice’ suggests that clear consent options are more likely following the introduction of guidance and standards that promote their implementation.

We observe three significant coefficient estimates for time effects (‘Wave Number’) across US and EU websites. Fist, we observe a positive significant result for the ‘Strong Nudge Consent’ response on US websites – suggesting that screenshots observed during later waves from US websites are generally more likely to contain strong nudges relative to no notice. Second, we observe negative significant coefficient estimates for ‘Cookie Banner’ and ‘Cookie Wall’ responses on EU websites. This suggests that as time progressed following GDPR enforcement, the likelihood of responding with either a cookie banner or cookie wall generally decreased. Each of the coefficient estimates show in part a potential movement over time towards responses that grant users options for consent.

Table 4.4: Multinomial Logit Model of All Responses: US Websites

		Dependent variable:				
				Response		
	Block EU Users	Cookie Banner	Cookie Wall	Full Consent	Strong Nudge Consent	Weak Nudge Consent
(Intercept)	−0.131	-1.598^{***}	-5.003^{***}	-3.529^{***}	-3.733^{***}	-3.346^{***}
	(0.317)	(0.226)	(0.856)	(0.477)	(0.386)	(0.444)
US IP	-22.313^{***}	-0.818^{***}	-5.174^{***}	-21.480^{***}	-2.810^{***}	-2.678^{***}
	(0.133)	(0.16)	(1.205)	(0.185)	(0.386)	(0.462)
Share EU Users	-49.026^{***}	5.910^{***}	7.471^{***}	7.786^{***}	10.656^{***}	11.932^{***}
	(9.332)	(1.539)	(2.534)	(2.327)	(2.052)	(1.953)
Reach per Million	-0.002^{**}	0.0001	0.001^{***}	0.0003	0.001^{***}	0.001^{***}
	(0.001)	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Events 1	0.22	0.455^{***}	−0.286	0.017	−0.207	0.351
	(0.17)	(0.143)	(0.504)	(0.284)	(0.227)	(0.266)
Events 2	0.0001	-0.494^{***}	-0.832^{**}	−0.151	-0.432^{***}	0.003
	(0.118)	(0.122)	(0.352)	(0.182)	(0.165)	(0.16)
Wave Number	−0.016	−0.029	0.085	0.069	0.125^{***}	0.022
	(0.03)	(0.026)	(0.089)	(0.052)	(0.044)	(0.05)
Cluster-Robust Errors	Yes	Yes	Yes	Yes	Yes	Yes

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.5: Multinomial Logit Model of All Responses: EU Websites

		Dependent variable:				
		Response				
	Block EU Users	Cookie Banner	Cookie Wall	Full Consent	Strong Nudge Consent	Weak Nudge Consent
(Intercept)	-	1.047***	−22.930***	−1.157	−0.783**	−0.263
		(0.307)	(3.926)	(0.791)	(0.36)	(0.317)
US IP	-	−0.565***	−0.487**	−1.239***	−0.840***	−0.761***
		(0.118)	(0.22)	(0.381)	(0.135)	(0.13)
Share EU Users	-	0.446	24.805***	−1.008	0.971***	0.044
		(0.312)	(4.201)	(0.682)	(0.363)	(0.308)
Reach per Million	-	0.0001	−0.009**	−0.0002	−0.0004	0.0001
		(0.0001)	(0.005)	(0.0005)	(0.0002)	(0.0001)
Events 1	-	0.305**	0.21	0.13	0.429***	0.076
		(0.121)	(0.158)	(0.427)	(0.135)	(0.125)
Events 2	-	0.13	0.235	0.594**	−0.083	0.459***
		(0.084)	(0.146)	(0.28)	(0.101)	(0.097)
Wave Number	-	−0.114***	−0.097***	−0.046	−0.011	0.015
		(0.02)	(0.028)	(0.062)	(0.026)	(0.022)
Cluster-Robust Errors	Yes	Yes	Yes	Yes	Yes	Yes
Note:						*p<0.1; **p<0.05; ***p<0.01

Table 4.6 and Table 4.7 show the coefficient estimates with robust standard errors clustered at the website-level for the subset of screenshots with consent notices from US and EU websites respectively. The coefficients in these models report the log-likelihood that a screenshot would adopt the given response relative to the 'Full Consent' response.

For US websites, the coefficients for 'Share EU Users' and 'Events 2' are significant for 'Weak Nudge Consent' and 'Strong Nudge Consent' respectively. The first of the coefficient estimates suggests that US websites with higher shares of EU users are more likely in general to implement a weak nudge relative to full consent with a decline button. The second coefficient suggests that the likelihood of strong nudges decreases relative to full consent following cookie guidance and the introduction of the IAB TCF v2.0.

For EU websites, the coefficients for 'Share EU Users' and 'Events 2' are significant, but of opposite signs. The coefficient for 'Share EU Users' suggest that strong nudges are more likely in general for EU websites for higher shares of EU users relative to full consent, while the coefficient for 'Events 2' suggests that strong nudges are less likely following the CNIL cookie guidance and the release of the IAB TCF v2.0 relative to full consent.

Table 4.6: Multinomial Logit Model of Consent Responses: US Websites

	Dependent variable:	
	Response	
	Strong Nudge Consent	Weak Nudge Consent
(Intercept)	−0.911	−0.604
	(0.636)	(0.673)
Share EU Users	5.540*	7.135**
	(2.977)	(2.926)
Reach per Million	0.002*	0.002*
	(0.001)	(0.001)
Events 1	−0.239	0.332
	(0.333)	(0.319)
Events 2	−0.371**	0.041
	(0.18)	(0.193)
Wave Number	0.094	−0.004
	(0.06)	(0.065)
Cluster-Robust Errors	Yes	Yes

Note: *p<0.1; **p<0.05; ***p<0.01

Table 4.7: Multinomial Logit Model of Consent Responses: EU Websites

	Dependent variable:	
	Response	
	Strong Nudge Consent	Weak Nudge Consent
(Intercept)	−0.062	0.458
	(1.038)	(1.028)
US IP	0.325	0.45
	(0.369)	(0.369)
Share EU Users	2.377***	1.404*
	(0.843)	(0.83)
Reach per Million	−0.00005	0.001
	(0.001)	(0.001)
Events 1	0.284	−0.100
	(0.445)	(0.433)
Events 2	−0.754**	−0.172
	(0.35)	(0.351)
Wave Number	0.05	0.074
	(0.07)	(0.069)
Cluster-Robust Errors	Yes	Yes
Note:		*p<0.1; **p<0.05; ***p<0.01

4.4.3 Logistic Regression

We use logistic regression to model how the website-level characteristics, traffic variables, and external events potentially impact the likelihood of different actions appearing within screenshots that contain consent notices. We narrow our focus to the consent sub-sample in this analysis as we expect actions such as ‘decline’ and ‘preference check boxes’ to only appear within consent notices. In these models, we examine seven DVs – one for each action in our screenshot rubric. This differs from of multinomial models in that each action is modeled separately. Our control variables are website characteristics including IP address of observation, share of EU users, and reach per million. We additionally estimate effects that may result from the wave during which the screenshot was observed.

Table 4.8: Logistic Regression Model of Consent Notice Actions: US Websites

	Dependent variable:						
	Accept	Decline	More Info	Settings	Close	Priv Pol	Pref Boxes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
US IP	-1.479^{***}	-17.481^{***}	0.461	0.414	1.516^{***}	-0.276	-1.978^{**}
	(0.484)	(0.359)	(0.550)	(0.478)	(0.465)	(0.511)	(0.979)
Share EU Users	0.085	-6.405^{*}	-0.138	2.045	-1.173	-0.170	-33.303^{***}
	(2.331)	(3.328)	(0.979)	(1.503)	(2.044)	(1.146)	(10.395)
Reach per Million	-0.0001	-0.002^{**}	0.0003	0.0002	0.0001	0.0002	-0.001
	(0.0002)	(0.001)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.001)
Events 1	-0.048	-0.069	-0.441	0.353	-0.847^{***}	0.191	0.620
	(0.490)	(0.305)	(0.295)	(0.241)	(0.282)	(0.273)	(0.553)
Events 2	0.221	0.155	-0.200	0.230	-0.327^{*}	0.071	-0.571
	(0.209)	(0.163)	(0.140)	(0.157)	(0.185)	(0.135)	(0.354)
Wave Number	-0.016	-0.037	0.111^{**}	-0.048	0.127^{**}	-0.045	-0.079
	(0.070)	(0.054)	(0.056)	(0.052)	(0.054)	(0.046)	(0.060)
Constant	2.732^{***}	0.192	-0.344	-0.408	-2.408^{***}	-0.160	-0.435
	(0.639)	(0.574)	(0.528)	(0.525)	(0.643)	(0.482)	(0.717)
Cluster-Robust Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 4.9: Logistic Regression Model of Consent Notice Actions: EU Websites

	Dependent variable:						
	Accept	Decline	More Info	Settings	Close	Priv Pol	Pref Boxes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
US IP	-0.229^{***}	-0.392^{*}	-0.172^{**}	0.135^{**}	0.184^{**}	0.222^{***}	0.438
	(0.074)	(0.237)	(0.069)	(0.065)	(0.071)	(0.072)	(0.323)
Share EU Users	-2.077	-1.820	-0.054	-0.557	1.941^{*}	0.106	-1.177
	(1.402)	(1.222)	(0.549)	(0.532)	(1.043)	(0.571)	(2.182)
Reach per Million	-0.00001	-0.0005	-0.0003	0.001^{**}	0.00005	0.0001	-0.005
	(0.0004)	(0.001)	(0.0002)	(0.001)	(0.0003)	(0.0003)	(0.007)
Events 1	0.198	-0.083	0.030	-0.298^{**}	-0.308	0.103	-11.319^{***}
	(0.292)	(0.481)	(0.179)	(0.152)	(0.246)	(0.179)	(3.004)
Events 2	-0.368^{**}	0.434	-0.159	0.525^{***}	0.291^{**}	-0.078	1.671^{***}
	(0.166)	(0.352)	(0.122)	(0.119)	(0.137)	(0.098)	(0.515)
Wave Number	0.170^{***}	-0.064	0.006	0.023	-0.135^{***}	0.008	-0.981^{**}
	(0.040)	(0.066)	(0.033)	(0.029)	(0.037)	(0.032)	(0.399)
Constant	2.072	-0.923	1.066^{*}	0.043	-1.757^{*}	-1.232^{**}	2.634
	(1.263)	(1.350)	(0.552)	(0.520)	(0.959)	(0.557)	(3.408)
Cluster-Robust Errors	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Note:							*p<0.1; **p<0.05; ***p<0.01

Table 4.8 and Table 4.9 show the coefficient estimates with cluster-robust standard errors for our logistic regression models. For many of the actions that may be present within a consent notice, observation IP is the most consistent predictor. When observed by a US visitor, US websites are less likely to provide users with actions that may make it easier for users to indicate consent: 'accept' buttons, 'decline' buttons, and cookie preference checkboxes ('Pref boxes' in tables). For each of these actions, the coefficient estimates are negative and significant. In contrast the coefficient estimate for the 'close' button is positive and significant, suggesting it is more likely to appear for US visitors. EU websites are also less likely to show US visitors 'accept' and 'decline' buttons within consent notices, but are more likely to display 'settings' buttons and links to privacy policies. Together, these coefficients may suggest that where US visitors are presented with consent notices, the consent options may be more difficult to navigate.

Among the traffic variables, 'Share EU Users' and 'Reach per Million' are negative and significant for preference checkboxes and decline buttons respectively for US websites. Likewise, the coefficient estimate for 'Reach per Million' on the likelihood of a settings button being present is statistically significant and positive for EU websites. These suggest that as traffic in general increases, the likelihood of showing a decline button within consent notices decreases for US websites and the likelihood of showing a settings button increases for EU websites. As the share of EU users increases, the likelihood of showing preference checkboxes, decreases for US websites.

The 'Events 1' coefficient has a significant negative impact on the likelihood of the 'close' button appearing within a consent notice for US websites. While this may be due to the CNIL Google fine (the Dutch DPA cookie wall guidance should not apply here since we only consider observations of consent notices), we might also expect to see impacts on other consent related actions if this were the case.

Both ‘Events 1’ and ‘Events 2’ coefficients appear to have stronger impacts for EU websites. For ‘Events 1’, we observe negative significant effects for ‘Settings’ buttons and ’Pref Boxes’. For ‘Events 2’, we observe a negative significant effect for ‘Accept’ buttons, and positive significant effects for ‘Settings’, ‘Close’, and ‘Pref Boxes’ buttons. Taken together, the changes in consent notice features predicted by these estimates may not be entirely consistent with what we might expect to observe following government and industry actions. While the net-estimate of for the ‘Settings’ button is positive over both events (suggesting an increase in clear consent options), the net-estimate for the ‘Action’ button is negative. Given that the CNIL guidance centered on eliciting clear consent from users, we might expect the likelihood of accept button appearing within consent notices to increase. While industry and government action may have encourage an increase in the adoption of consent notices, it does not seem to have impacted their features in a consistent manner.

The ‘Wave Number’ coefficient has positive significant estimates for ‘More Info’ and ‘Close’ buttons on US websites, suggesting that the likelihood of these actions appearing within a consent notice increases over the time period of our study. In contrast, the ‘Wave Number’ coefficient estimate for the ‘Accept’ button is negative and significant while the estimates for the ‘Close’ and ‘Pref Boxes’ options are negative and significant. These suggest that the likelihood of an accept button appearing within a consent notice for a EU website increases over time while the likelihood that close buttons or preference checkboxes appearing decrease over time. Taken together, the results do not paint a clear picture regarding a consistent shift in consent notice design over time. While some coefficient estimates (such as the ‘Accept’ button on EU websites) may suggest increased adoption in the types of actions needed to provided clear consent, these effects are not consistent across US and EU websites. Only in the case of the ‘Close’ button are the coefficient estimates significant for both US and EU websites. Although both significant, the directions of the coefficients are in opposite directions, suggesting a divergence over time – with the likelihood of the button appearing increasing over time for US websites while decreasing for EU websites. The behavioral implications of this are mixed. Because close buttons are often used to allow users to dismiss consent notices without taking further action, they are potentially incompatible with consent guidance (such as that issued by the CNIL) that requires consent to be clear and unambiguous. For EU websites, a decrease in 'close' buttons over time could be interpreted as a shift towards adherence with these guidelines. For US websites, an increase in close buttons (along with an increase in more information buttons) may suggest an increase in consent notice designs that attempt to make accepting tracking easier than rejecting tracking.

4.5 Discussion

Overall, our results highlight a geographic disparity in website response and consent notice design between the US and EU after the GDPR. In our multinomial logit models, we find that US visitors to US and EU websites are less likely to receive any type of privacy messaging relative to EU visitors. Relatedly, we find that US websites are more likely in general to provide privacy messaging to visitors if the share of those visitors originating from the EU is higher. Our logistic regression models of consent notice features provide additional detail to these results, finding that US websites are generally less likely to provide users with consent options such as 'accept' and 'decline' buttons while being more likely to show 'close' buttons. Because the default behavior of many websites is to track users by default, an increased prevalence 'close' buttons may result in users dismissing these notices without providing clear and unambiguous consent [86]. These findings are consistent for EU websites, with the exceptions that the coefficient for the 'decline' button is only significant at the $p < 0.1$ level and the likelihood of the inclusion of a 'settings' button within consent notices is higher for US visitors relative to EU visitors. While the difference in response selection among websites based on visitor location may be expected based on the potential motivation of many website operators to maximize targeted ad revenue, the differences in the features of consent notices based on visitor location is more surprising.

Our analyses of events following the GDPR find evidence that government and industry actions had an impact on types of responses adopted by websites and the features present within consent notices. Cumulatively, our multinomial models for all responses across US and EU websites suggest a move away from responses that make it difficult for users to reject tracking (on US websites) and towards consent options that are more clear and easier for users to navigate (on EU websites). Notably, the effects of 'Events 2' appear to be stronger than 'Events 1', where we observe more significant coefficient estimates. This could suggest that the CNIL cookie guidance and the introduction of IAB's TCF v2.0 were stronger events in moving response selection compared to the CNIL's €50 million fine against Google and the Dutch DPA's cookie wall guidance. These overall findings are further supported by our multinomial model of consent notices, which finds that 'Events 2' reduced the occurrence of 'Strong Nudge Consent' responses relative to 'Full Consent'. In contrast, the results of our analysis of consent notice features are not always consistent with what we might expect following government and industry actions. Overall, the events we consider in this study appear to potentially move websites away from responses that make it hard for users to reject tracking and towards consent responses, but do not impact the features present within consent notices in a manner consistent with government guidance.

Importantly, our ability to make claims about individual enforcement actions or guidance is limited due to the granularity of our data, where the size of gaps between waves towards the end of our data collection period make it so we cannot claim with certainty that the observed effects are due to any one event. Future studies seeking to examine the impacts of enforcement and guidance actions may seek to gather more granular data and to identify the effects of individual events with more specificity.

While our study is unique in its ability to examine the factors which impact websites' GDPR response over time, our screenshot data limits the depth of analysis we are able to perform. Specifically, we only able to examine the 'surface-level' of privacy notices that are visible within screenshots. We are unable to examine the designs and features of secondary screens or pop-ups which may appear after clicking on buttons within the notice (such as 'settings'). Future studies may wish to examine the dynamics of these secondary screens.

Throughout our study, we make several assumptions which may have impacted our results. We distinguish between nudges within consent notices based on the presence of a 'settings' button. It is possible that different definitions of 'nudges' may yield slightly different coefficient estimates. Additionally, we rely on qualitative coding to identify consent mechanisms. Specifically, we flagged notices which 'implied' the ability to decline tracking as providing some option for consent. While we verified the accuracy of this coding by manually examining multiple random samples of screenshots labeled as containing a consent notice, some bias regarding definitions of consent is potentially present resulting from our coding process. However, we expect this bias to be minimal and consistent across our data.

The results of our analysis open up several possible directions for future research that would continue to build upon the understanding of how websites responded to the GDPR. While our analysis focused on 'News and Media' websites – as we in part expect these websites to be more likely to implement consent notices – future work may seek to include websites from other categories. Additionally, future work may wish to incorporate data on Consent Management Platforms to identify how their adoption over time impacts the response choice of websites and the features within consent notices over time. As part of our OpenWPM data, we gathered all cookies placed by each of the websites we visited. It may be possible to identify the presence of CMPs using these cookies.

4.6 Conclusion

We examined $14,832$ screenshots from a longitudinal panel of 911 'News and Media' websites to identify trends in privacy notice design following the enforcement of the GDPR in May 2018. Specifically, we examine whether website-level characteristics such as traffic and visitor location impact the types of notices websites use over time and the features present within consent notices over time. We additionally study the potential impact of four government and industry actions which may have impacted the types and features of notices following their occurrence. To facilitate our analysis, we developed a rubric of privacy notice features and manually assigned each screenshot to one of seven possible 'responses' that websites might take in response to the GDPR. We use multinomial logistic regression to examine the adoption of different response types and logistic regression to examine changes in the features of consent notices. Our results identify a geographic disparity in privacy messaging, where US visitors to US and EU websites are less likely to receive privacy messaging relative to EU visitors. We additionally find that screenshots on US websites are more likely to contain privacy notices if the website receives a large share of its traffic from the EU. Our events analysis finds that government and industry actions generally reduced the likelihood of US websites adopting responses that make it more difficult for users to reject tracking, while increasing the likelihood of EU websites adopting responses that make it easier for users to reject tracking. While our coefficient estimates suggest that the events may have helped push websites towards adopting clearer consent notices in line with guidance issued by European regulators, the granularity and specificity of our data make it difficult to draw any causal relationships from any single action. Future studies may seek to build on our work by incorporating data on Consent Management Platforms, secondary consent screens, and websites from categories outside 'News and Media'.

Chapter 5

Conclusion

In the previous three chapters, I investigated different behavioral interventions for privacy decisions across three contexts. In Chapter 2, I examined the idea of 'tailoring' privacy nudges to users based on their decision-making and personality traits. Over three studies (and 12+ individual experiments), I examined whether the differences in disclosure rates elicited by placing 'framing' and 'social norms' nudges within disclosure decisions could be predicted by different decision making and personality traits. While 'Study 2' did identify potential evidence of these tailoring effects, our followup replication in 'Study 3' did not confirm their existence – suggesting that the effects of tailoring nudges for privacy disclosure decisions are fragile if present.

Where Chapter 2 sought to improve upon methods for behavioral interventions, in Chapter 3 I sought to apply an existing method for behavioral interventions to the new context of mobile privacy decisions. Specifically, I investigated whether three different styles of explanations ('behavioral', 'social', and 'hybrid') impacted the likelihood that participants would choose to adopt hypothetical recommendations for changes to mobile permission settings. The results from an online survey of participants did not find significant effects for any of the tested explanation styles, both when measured relative to participants that received no explanation, and when measured relative to each other. Instead, I found the control variables including perceived recommendation quality and general online privacy preferences to be much stronger predictors of adoption intent. A followup analysis of participants' qualitative responses found that participants' questioned the purpose provided for the hypothetical permission request within the disclosure scenario. Together, these findings suggest a limit to the effectiveness of behavioral interventions in the face of pre-existing privacy preferences.

Finally, in Chapter 4 I explored the presence of different behavioral interventions within privacy notices on 911 'News and Meida' websites following the enforcement of the GDPR in May 2018. I developed a rubric for privacy notices which was used to classify $14,832$ screenshots according to the features present within privacy notices. I further used these features to develop 'response' labels for how websites responded to the GDPR and identified four government and industry actions that may have impacted websites' responses. Across multiple regression models, I identified a geographic disparity in the presence of privacy messaging on websites – where US visitors were less likely to receive privacy messaging relative to EU visitors. I additionally identified evidence of effects resulting from industry and government actions on websites' response choice, where the likelihood of US websites responding with notices that make it difficult or impossible for visitors to reject tracking decreased while the likelihood of EU websites responding with consent notices that make it easier for visitors to reject tracking increased. Together, these results suggest that while the GDPR initially did not lead to an increase in clear consent options for users, later actions have gradually moved websites towards consent responses.

5.1 Challenges of Behavioral Interventions and Recommendations

The three studies presented in this dissertation provide an insight into the challenges and limitations of behavioral interventions for privacy decisions and suggest lessons for practitioners. Many of the challenges of developing effective interventions arise from the nature of privacy decisions, which center on privacy preferences that are often highly contextual and can be poorly defined [5]. This makes behavioral interventions such as nudges unique when applied to privacy decisions in contrast to domains such as security – where preferences are not expected to factor into decisions and an objectively preferred outcome (in terms of security) can often be identified. Choice architects should take care to properly consider the role of privacy preferences when designing these types of interventions.

Privacy preferences impact both the effectiveness of behavioral interventions and the ethics around their deployment. As observed in both Chapter 2 and Chapter 3, the effects of nudges and explanations (arguably a type of nudge itself) are sensitive to strong preferences. In Chapter 2, we iterated on the design of our disclosure scenario to arrive at one where the aggregate disclosure rate was roughly equal. We found these scenarios to elicit the strongest differences in disclosure rates, which were essential to testing our main hypotheses. In contrast, the hypothetical disclosure scenario in Chapter 3 produced lopsided disclosure rates for participants that received a deny recommendation (where most participants chose to deny the hypothetical disclosure), likely contributing to our null findings. While behavioral interventions can be very effective, they may operate best at the margins where privacy preferences are more uncertain.

Finally, nudges necessarily promote a specific outcome by modifying the structure and presentation of decisions. In the context of privacy decisions, this necessitates knowing a user's privacy preferences. Because the complexity of privacy preferences often make it difficult to predict a user's privacy preferences in a particular context, the use of nudges in privacy decisions may lead to negative outcomes for users. In many cases, behavioral interventions such as explanations which aid decision making by providing additional information or context may be preferable. When seeking to use a nudge, choice architects should first ensure that users' privacy preferences are known.

5.2 Policy Implications

An understanding of how users make privacy decisions will be essential to the crafting of successful consumer privacy regulation in the US. At the time of this writing (August 2022), five states (California, Colorado, Connecticut, Utah, and Virginia as of June 2022) have passed legislation on consumer privacy. Many of these laws contain requirements that parties which collect and process consumer data provide the ability for users to opt-out. As observed in Chapter 2 and Chapter 3 the design of these disclosure choices is important, where defaults and framing can have powerful effects on user behavior. As observed in Chapter 4, the design of privacy notices was inconsistent immediately after the introduction of the GDPR. Many sites which implemented consent notices often did so with nudges that would make rejecting tracking more difficult for users. Future consumer privacy regulation should consider introducing enforceable design guidelines alongside regulation to reduce the likelihood of these 'dark patterns'.

However, where possible, consumer privacy regulation should seek to reduce the over-reliance on ‘notice and choice’ as a tool of consumer privacy protection. Because of data sharing with third-parties and increasingly sophisticated methods for drawing new inferences from existing data, it can be difficult or impossible for users to assess the privacy risks of disclosures. Reducing the number of instances where ‘notice and choice’ is misapplied may ultimately reduce the need for behavioral interventions which seek to help users overcome hurdles such as information asymmetries, if only partially. Where ‘notice and choice’ is applied, building more protections around data usage rather than data collection may help reduce cognitive hurdles for users. This is supported by the findings in the followup analysis in Chapter 3, where one of participants’ primary concerns when considering the hypothetical recommendation was the necessity of the purpose given for requesting mobile location data. Framing consent choices around the purposes for data usage may therefore help users more easily protect the contextual integrity of information flows.

Bibliography

[1] A. Acquisti, L. Brandimarte, and G. Loewenstein. "Privacy and human behavior in the age of information". In: *Science* 347 (6221 2015), pp. 509–514. DOI: 10.1126/science.aaa1465.

[2] A Acquisti, LK John, and G Loewenstein. "The impact of relative standards on the propensity to disclose". In: *Journal of Marketing Research* 49.2 (2012), pp. 160–174. ISSN: 0022-2437. DOI: 10.1509/jmr.09.0215. URL: http://journals.ama.org/doi/abs/10.1509/jmr.09.0215.

[3] Alessandro Acquisti. "Nudging privacy: The behavioral economics of personal information". In: *IEEE Security Privacy* 7 (6 2009), pp. 82–85. ISSN: 15407993. DOI: 10.3233/978-1-61499-057-4-193.

[4] Alessandro Acquisti et al. "Nudges for Privacy and Security : Understanding and Assisting Users ' Choices Online". In: (2016), pp. 1–40.

[5] Idris Adjerid, Alessandro Acquisti, and George Loewenstein. "Framing and the Malleability of Privacy Choices". In: (2014).

[6] Idris Adjerid, Eyal Peer, and Alessandro Acquisti. "Beyond the Privacy Paradox: Objective versus Relative Risk in Privacy Decision Making". 2016. DOI: 10.2139/ssrn.2765097.

[7] G. Adomavicius and A. Tuzhilin. "Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions". In: *IEEE Transactions on Knowledge and Data Engineering* 17 (6 June 2005), pp. 734–749. ISSN: 1041-4347. DOI: 10.1109/TKDE.2005.99.

[8] Hunt Allcott and Judd Kessler. "The Welfare Effects of Nudges : A case study of energy use social comparisons". In: (2015).

[9] Hazim Almuhimedi et al. "Your Location has been Shared 5,398 Times! A Field Study on Mobile App Privacy Nudging". In: (2014).

[10] Theo Araujo et al. "In AI we trust? Perceptions about automated decision-making by artificial intelligence". In: *AI SOCIETY* 35 (3 Sept. 2020), pp. 611–623. ISSN: 0951-5666. DOI: 10.1007/s00146-019-00931-w.

[11] Rebecca Balebako et al. "Nudging Users Towards Privacy on Mobile Devices". In: *CEUR Workshop Proceedings* 722 (2011), pp. 23–26. ISSN: 16130073.

[12] Murray R Barrick and Michael K Mount. “The Big Five Personality Dimensions and Job Performance : A Meta Analysis”. In: *Personnel Psychology* 44.1 (1991), pp. 1–26. ISSN: 0031-5826. DOI: 10.1111/j.1744-6570.1991.tb00688.x.

[13] M. Bashir et al. “Online privacy and informed consent: The dilemma of information asymmetry”. In: *Proceedings of the Association for Information Science and Technology* 52 (1 2015), pp. 1–10. DOI: 10.1002/pra2.2015.145052010043.

[14] John Beshears et al. “Who Is Easier to Nudge?” In: *NBER Working Paper* 401 (2015). DOI: 10.1007/s13398-014-0173-7.2.

[15] Daniel Billsus and Michael J. Pazzani. “A personal news agent that talks, learns and explains”. In: ACM Press, 1999, pp. 268–275. ISBN: 158113066X. DOI: 10.1145/301136.301208.

[16] European Data Protection Board. *The CNIL’s restricted committee imposes a financial penalty of 50 Million euros against GOOGLE LLC.* Jan. 2019. URL: https://edpb.europa.eu/news/national-news/2019/cnils-restricted-committee-imposes-financial-penalty-50-million-euros_en.

[17] Wändi Bruine de Bruin, Andrew M Parker, and Baruch Fischhoff. “Adult Decision-Making Competence Inventory”. In: (2007), pp. 938–956.

[18] Wändi Bruine de Bruin, Andrew M Parker, and Baruch Fischhoff. “Individual Differences in Adult Decision-Making Competence.” In: *Journal of Personality and Social Psychology* 92.5 (2007), pp. 938–956. ISSN: 0022-3514. DOI: 10.1037/0022-3514.92.5.938.

[19] J T Cacioppo, R E Petty, and C F Kao. *The efficient assessment of need for cognition.* 1984. DOI: 10.1207/s15327752jpa4803_13.

[20] R Casadesus-Masanell and F Zhu. “Business Model Innovation and Competitive Imitation: The Case of Sponsor-Based Business Models”. In: *Strategic Management Journal* 34.4 (2013), pp. 464–482.

[21] Daniel Castro. *Stricter Privacy Regulations for Online Advertising Will Harm the Free Internet.* Technical Report. Information Technology and Innovation Foundation, 2010. URL: https://itif.org/publications/2010/09/08/stricter-privacy-regulations-online-advertising-will-harm-free-internet#:~:text=Stricter\%20Privacy\%20Regulations\%20for\%20Online\%20Advertising\%20Will\%20Harm\%20the\%20Free\%20Internet,-Daniel\%20Castro\%20September&text=A\%20study\%20shows\%20that\%20overly,effectiveness\%20of\%20the\%20Internet\%20ecosystem..

[22] Long-Sheng Chen et al. “Developing recommender systems with the consideration of product profitability for sellers”. In: *Information Sciences* 178 (4 Feb. 2008), pp. 1032–1048. ISSN: 00200255. DOI: 10.1016/j.ins.2007.09.027.

[23] Lynne M. Coventry et al. "Personality and social framing in privacy decision-making: A study on cookie acceptance". In: *Frontiers in Psychology* 7.SEP (2016), pp. 1–12. ISSN: 16641078. DOI: 10.3389/fpsyg.2016.01341.

[24] Henriette Cramer et al. "The effects of transparency on trust in and acceptance of a content-based art recommender". In: *User Modeling and User-Adapted Interaction* 18 (5 Nov. 2008), pp. 455–496. ISSN: 09241868. DOI: 10.1007/s11257-008-9051-3.

[25] Yves Croissant. "Estimation of Random Utility Models in R: The mlogit Package". In: *Journal of Statistical Software* 95 (11 2020). ISSN: 1548-7660. DOI: 10.18637/jss.v095.i11.

[26] Martin Degeling et al. "We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy". In: 2019.

[27] Gregor Dorfleitner, Lars Hornuf, and Julia Kreppmeier. "Promise Not Fulfilled: Fintech Data Privacy, and the GDPR". In: *SSRN Electronic Journal* (2021). ISSN: 1556-5068. DOI: 10.2139/ssrn.3950094.

[28] James N Druckman et al. *On the Limits of Framing Effects: Who Can Frame?* Tech. rep. 4. 2001, p. 18. URL: http://www.journals.uchicago.edu/t-and-c.

[29] Caitlin Drummond and Baruch Fischhoff. "Development and Validation of the Scientific Reasoning Scale". In: *Journal of Behavioral Decision Making* 30.1 (2017), pp. 26–38. ISSN: 10990771. DOI: 10.1002/bdm.1906.

[30] Serge Egelman and Eyal Peer. "The Myth of the Average User". In: *NSPW 2015 Proceedings* (2015), pp. 16–28.

[31] Rob van Eijk et al. "The impact of user location on cookie notices (inside and outside of the European union)". In: 2019.

[32] IAB Europe. *GDPR Guidance: Legitimate Interests Assessments (LIA) for Digital Advertising.* Mar. 2021.

[33] Interactive Advertising Bureau Europe. *IAB Europe IAB Tech Lab release updated Transparency Consent Framework.* Aug. 2019. URL: https://iabeurope.eu/press-releases/iab-europe-iab-tech-lab-release-updated-transparency-consent-framework/.

[34] Interactive Advertising Bureau Europe. *Transparency and Consent Framework v2.0.* Aug. 2019. URL: https://iabeurope.eu/tcf-2-0/.

[35] Alan Fernihough. *mfx: Marginal Effects, Odds Ratios and Incidence Rate Ratios for GLMs.* 2014. URL: https://cran.r-project.org/package=mfx.

[36] John Fox and Sanford Weisberg. *An {R} Companion to Applied Regression.* Second. Thousand Oaks {CA}: Sage, 2011. URL: http://socserv.socsci.mcmaster.ca/jfox/Books/Companion.

[37] A Goldfarb. "Concentration in Advertising-Supported Online Markets: An Empirical Approach". In: *Economics of Innovation and New Technology* 13.6 (2004), pp. 581–594.

[38] Daniel G. Goldstein et al. "Beyond nudges: Tools of a choice architecture". In: *Marketing Letters* 23.2 (2012), pp. 487–504. ISSN: 0923-0645. DOI: 10.1007/s11002-012-9186-1.

[39] Asela Gunawardana and Guy Shani. *Evaluating Recommender Systems*. 2015. DOI: 10.1007/978-1-4899-7637-6_8. URL: https://link.springer.com/chapter/10.1007/978-1-4899-7637-6_8.

[40] David Gunning and David Aha. "DARPA's Explainable Artificial Intelligence (XAI) Program". In: *AI Magazine* 40 (2 June 2019). ISSN: 2371-9621. DOI: 10.1609/aimag.v40i2.2850.

[41] Hana Habib et al. ""It's a scavenger hunt": Usability of Websites' Opt-Out and Data Deletion Choices". In: ACM, Apr. 2020. ISBN: 9781450367080. DOI: 10.1145/3313831.3376511.

[42] Hana Habib et al. "“Okay, whatever": An Evaluation of Cookie Consent Interfaces". In: ACM, Apr. 2022, pp. 1–27. ISBN: 9781450391573. DOI: 10.1145/3491102.3501985.

[43] Andrew F. Hayes and Jörg Matthes. "Computational procedures for probing interactions in OLS and logistic regression: SPSS and SAS implementations". In: *Behavior Research Methods* 41.3 (2009), pp. 924–936. ISSN: 1554351X. DOI: 10.3758/BRM.41.3.924.

[44] Jonathan L. Herlocker, Joseph A. Konstan, and John Riedl. "Explaining collaborative filtering recommendations". In: ACM Press, 2000, pp. 241–250. ISBN: 1581132220. DOI: 10.1145/358916.358995.

[45] Maximilian Hils, Daniel W. Woods, and Rainer Böhme. "Measuring the Emergence of Consent Management on the Web". In: ACM, Oct. 2020. ISBN: 9781450381383. DOI: 10.1145/3419394.3423647.

[46] Marek Hlavac. *stargazer: Well-Formatted Regression and Summary Statistics Tables*. Cambridge, USA, 2015. URL: http://cran.r-project.org/package=stargazer.

[47] IHS Technology. *Paving the Way: How Online Advertising Enables the Digital Economy of the Future*. Technical Report. 2015. URL: https://www.iabfrance.com/sites/www.iabfrance.com/files/atoms/files/iab_ihs_euro_ad_macro_finalpdf.pdf.

[48] Anthony Jameson et al. *Human Decision Making and Recommender Systems*. 2015. DOI: 10.1007/978-1-4899-7637-6_18.

[49] Yucheng Jin, Nava Tintarev, and Katrien Verbert. “Effects of personal characteristics on music recommender systems with different levels of controllability”. In: ACM, Sept. 2018. ISBN: 9781450359016. DOI: 10.1145/3240323.3240358. URL: https://dl.acm.org/doi/10.1145/3240323.3240358.

[50] Daniel Kahneman and Amos Tversky. “Prospect Theory: An Analysis of Decision under Risk”. In: *Source: Econometrica* 47.2 (1979), pp. 263–292. URL: http://www.jstor.org/stable/1914185http://www.jstor.org/stable/1914185?seq=1{\&}cid=pdf-reference{\#}references{_}tab{_}contentshttp://about.jstor.org/terms.

[51] Carmen Keller and Michael Siegrist. “Effect of risk communication formats on risk perception depending on numeracy”. In: *Medical Decision Making* 29.4 (2009), pp. 483–490. ISSN: 1552681X. DOI: 10.1177/0272989X09333122.

[52] Bart P. Knijnenburg and Alfred Kobsa. “Helping users with information disclosure decisions”. In: Association for Computing Machinery (ACM), 2013, p. 407. DOI: 10.1145/2449396.2449448.

[53] Johannes Kunkel et al. “Let Me Explain: Impact of Personal and Impersonal Explanations on Trust in Recommender Systems”. In: ACM, May 2019, pp. 1–12. ISBN: 9781450359702. DOI: 10.1145/3290605.3300717.

[54] Anton Kühberger, Michael Schulte-Mecklenbeck, and Josef Perner. “Framing decisions: Hypothetical and real”. In: *Organizational Behavior and Human Decision Processes* 89 (2 Nov. 2002), pp. 1162–1175. ISSN: 07495978. DOI: 10.1016/S0749-5978(02)00021-3.

[55] T B Lalitha and P S Sreeja. “Personalised Self-Directed Learning Recommendation System”. In: *Procedia Computer Science* 171 (2020), pp. 583–592. ISSN: 18770509. DOI: 10.1016/j.procs.2020.04.063.

[56] Anja Lambrecht et al. “How Do Firms Make Money Selling Digital Goods Online?” In: *Marketing Letters* 25.3 (2014), pp. 331–341.

[57] Vincent Lefrere et al. *The Impact of GDPR on Ad-Supported Content Providers*. 2020.

[58] Xenofon Liapakis. “A GDPR Implementation Guide for the Insurance Industry”. In: *International Journal of Reliable and Quality E-Healthcare* 7 (4 Oct. 2018), pp. 34–44. ISSN: 2160-9551. DOI: 10.4018/IJRQEH.2018100103.

[59] Commission nationale de l’informatique et des libertés. “Cookies and other tracking devices: the CNIL publishes new guidelines”. In: (July 2019). URL: https://www.cnil.fr/en/cookies-and-other-tracking-devices-cnil-publishes-new-guidelines.

[60] Isaac M Lipkus, Greg Samsa, and Barbara K Rimer. “General Performance on a Numeracy Scale among Highly Educated Samples”. In: 21 (2001), pp. 37–44. DOI: 10.1177/0272989X0102100105.

[61] Bin Liu et al. "Follow My Recommendations: A Personalized Privacy Assistant for Mobile App Permissions". In: *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)* (Soups 2016), pp. 27–41. URL: `https://www.usenix.org/conference/soups2016/technical-sessions/presentation/liu`.

[62] Rui Liu et al. "When Privacy Meets Usability: Unobtrusive Privacy Permission Recommendation System for Mobile Apps based on Crowdsourcing". In: *IEEE Transactions on Services Computing* (2016). ISSN: 1939-1374. DOI: `10.1109/TSC.2016.2605089`.

[63] Xueming Luo and Jie Zhang. "How Do Consumer Buzz and Traffic in Social Media Marketing Predict the Value of the Firm?" In: *Journal of Management Information Systems* 30.2 (Oct. 2013), pp. 213–238.

[64] Dominique Machuletz and Rainer Böhme. "Multiple Purposes, Multiple Problems: A User Study of Consent Dialogs after GDPR". In: *Proceedings on Privacy Enhancing Technologies* 2020 (2 Apr. 2020). ISSN: 2299-0984. DOI: `10.2478/popets-2020-0037`.

[65] Naresh K Malhotra, Sung S Kim, and James Agarwal. "Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model". In: *Information systems research* 15.4 (2004), pp. 336–355.

[66] Nathan Malkin et al. "Personalized Security Messaging: Nudges for Compliance with Browser Warnings". In: *EuroUSEC 17* April (2017). DOI: `10.14722/eurousec.2017.23008`. URL: `https://www.internetsociety.org/sites/default/files/eurousec2017{\_}08{\_}Malkin{\_}paper.pdf`.

[67] André Marchand and Paul Marx. "Automated Product Recommendations with Preference-Based Explanations". In: *Journal of Retailing* 96 (3 Sept. 2020), pp. 328–343. ISSN: 00224359. DOI: `10.1016/j.jretai.2020.01.001`.

[68] Estelle Massé. *Three Years Under the EU GDPR: An Implementation Progress Report.* Access Now, May 2021. URL: `https://www.accessnow.org/cms/assets/uploads/2021/05/Three-Years-Under-GDPR-report.pdf`.

[69] D. Harrison McKnight, Vivek Choudhury, and Charles Kacmar. "Developing and Validating Trust Measures for e-Commerce: An Integrative Typology". In: *Information Systems Research* 13 (3 Sept. 2002). ISSN: 1047-7047. DOI: `10.1287/isre.13.3.334.81`.

[70] D. Harrison McKnight, Larry L. Cummings, and Norman L. Chervany. "Initial Trust Formation in New Organizational Relationships". In: *Academy of Management Review* 23 (3 July 1998). ISSN: 0363-7425. DOI: `10.5465/amr.1998.926622`.

[71] Brent Mittelstadt, Chris Russell, and Sandra Wachter. "Explaining Explanations in AI". In: ACM, Jan. 2019, pp. 279–288. ISBN: 9781450361255. DOI: `10.1145/3287560.3287574`.

[72] Barrington Moore. *Privacy: Studies in Social and Cultural History*. 1984. ISBN: 9781138884700.

[73] Pardis Emami Naeini et al. “Privacy Expectations and Preferences in an IoT World”. In: USENIX Association, July 2017, pp. 399–412. ISBN: 978-1-931971-39-3. URL: `https://www.usenix.org/conference/soups2017/technical-sessions/presentation/naeini`.

[74] Helen Nissenbaum. “A Contextual Approach to Privacy Online”. In: *Daedalus* 140 (4 Oct. 2011), pp. 32–48. ISSN: 0011-5266. DOI: `10.1162/DAED_a_00113`.

[75] P A Norberg and D R Horne. “The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors - NORBERG - 2007 - Journal of Consumer Affairs - Wiley Online Library”. In: *Journal of Consumer ...* 41.1 (2007), pp. 100–126. ISSN: 00220078. DOI: `10.1111/j.1745-6606.2006.00070.x`. URL: `http://onlinelibrary.wiley.com/doi/10.1111/j.1745-6606.2006.00070.x/full{\%}5Cnpapers2://publication/uuid/99384401-B176-4BF1-9E42-A34D1BEFBCA5`.

[76] Midas Nouwens et al. “Dark Patterns after the GDPR: Scraping Consent Pop-ups and Demonstrating their Influence”. In: ACM, Apr. 2020. ISBN: 9781450367080. DOI: `10.1145/3313831.3376321`.

[77] Ingrid Nunes and Dietmar Jannach. “A systematic review and taxonomy of explanations in decision support and recommender systems”. In: *User Modeling and User-Adapted Interaction* 27 (3-5 Dec. 2017), pp. 393–444. ISSN: 0924-1868. DOI: `10.1007/s11257-017-9195-0`.

[78] Gabriele Paolacci and Jesse Chandler. “Inside the Turk: Understanding Mechanical Turk as a Participant Pool”. In: *Current Directions in Psychological Science* 23.3 (2014), pp. 184–188. ISSN: 14678721. DOI: `10.1177/0963721414531598`.

[79] Alexis Papadimitriou, Panagiotis Symeonidis, and Yannis Manolopoulos. “A generalized taxonomy of explanations styles for traditional and social recommender systems”. In: *Data Mining and Knowledge Discovery* 24 (3 May 2012). ISSN: 1384-5810. DOI: `10.1007/s10618-011-0215-0`. URL: `https://link.springer.com/article/10.1007/s10618-011-0215-0`.

[80] Yong Jin Park. “Digital Literacy and Privacy Behavior Online”. In: *Communication Research* 40.2 (2013), pp. 215–236. ISSN: 00936502. DOI: `10.1177/0093650211418338`.

[81] Eyal Peer et al. “Beyond the Turk: Alternative platforms for crowdsourcing behavioral research”. In: ().

[82] Eyal Peer et al. “Nudge Me Right: Personalizing Online Nudges to People’s Decision-Making Styles”. In: *SSRN Electronic Journal* (2019), pp. 1–23. DOI: `10.2139/ssrn.3324907`.

[83] Autoriteit Persoonsgegevens. *Websites moeten toegankelijk blijven bij weigeren tracking cookies*. Mar. 2019.

[84] Ellen Peters. “Numeracy and the perception and communication of risk”. In: *Annals of the New York Academy of Sciences* 1128 (2008), pp. 1–7. ISSN: 17496632. DOI: `10.1196/annals.1399.001`.

[85] Sonam Samat et al. “Format vs. Content: The Impact of Risk and Presentation on Disclosure Decisions”. In: *Symposium on Usable Privacy and Security (SOUPS) 2017* Soups (2017), p. 9. URL: `https://www.usenix.org/conference/soups2017/technical-sessions/presentation/samat-disclosure`.

[86] Iskander Sanchez-Rola et al. “Can I Opt Out Yet? GDPR and the Global Illusion of Cookie Control”. In: 2019, pp. 340–351.

[87] Florian Schaub, Rebecca Balebako, and Lorrie Faith Cranor. “Designing Effective Privacy Notices and Controls”. In: *IEEE Internet Computing* 21 (3 May 2017), pp. 70–77. ISSN: 1089-7801. DOI: `10.1109/MIC.2017.75`.

[88] Susanne Scott and Bruce Reginald. “Decision-Making Style: The Development and Assessment of a New Measure”. In: *Educational and psychological measurement* (1995), pp. 818–831.

[89] Susanne G. Scott and Reginald A. Bruce. “Decision-Making Style: The Development and Assessment of a New Measure”. In: *Educational and Psychological Measurement* 55.5 (1995), pp. 818–831. ISSN: 15523888. DOI: `10.1177/0013164495055005017`. arXiv: `0803973233`.

[90] Amit Sharma and Dan Cosley. “Do social explanations work?: studying and modeling the effects of social explanations in recommender systems”. In: ACM Press, 2013, pp. 1133–1144. ISBN: 9781450320351. DOI: `10.1145/2488388.2488487`.

[91] Benjamin Shiller, Joel Waldfogel, and Johnny Ryan. “The Effect of Ad Blocking on Website Traffic and Quality”. In: *The RAND Journal of Economics* 49.1 (2018), pp. 43–63.

[92] J.H Smith, S.J Milberg, and S.J Burke. “Information privacy: measuring individuals concerns about organizational practices.” In: *MIS quaterly* 20.2 (1996), pp. 167–196.

[93] D.J. Solove. “Privacy self-management and the consent dilemma”. In: *Harvard Law Review* 126 (7 2013), pp. 1880–1903.

[94] Jannick Sørensen and Sokol Kosta. “Before and After GDPR: The Changes in Third Party Presence at Public and Private European Websites”. In: *The World Wide Web Conference*. 2019, pp. 1590–1600.

[95] Neil Stewart et al. “The average laboratory samples a population of 7,300 Amazon Mechanical Turk workers”. In: *Judgment and Decision Making* 10.5 (2015), pp. 479–491. ISSN: 1930-2975. DOI: 10.1017/CBO9781107415324.004. arXiv: arXiv:1011.1669v3.

[96] Cass R. Sunstein. “Nudging: A Very Short Guide”. In: *Journal of Consumer Policy* 37.4 (2014), pp. 583–588. ISSN: 15730700. DOI: 10.1007/s10603-014-9273-1. arXiv: arXiv:1011.1669v3.

[97] Martin Svrcek, Michal Kompan, and Maria Bielikova. “Towards understandable personalized recommendations: Hybrid explanations”. In: *Computer Science and Information Systems* 16 (1 2019). ISSN: 1820-0214. DOI: 10.2298/CSIS171217012S. URL: http://www.doiserbia.nb.rs/img/doi/1820-0214/2019/1820-02141800012S.pdf.

[98] Richard Thaler and Cass Sunstein. *Nudge: Improving Decisions About health, Wealth, and Happiness.* Penguin Books, 2009.

[99] Nava Tintarev and Judith Masthoff. “Evaluating the effectiveness of explanations for recommender systems”. In: *User Modeling and User-Adapted Interaction* 22 (4-5 Oct. 2012), pp. 399–439. ISSN: 0924-1868. DOI: 10.1007/s11257-011-9117-5.

[100] Nava Tintarev and Judith Masthoff. *Explaining Recommendations: Design and Evaluation.* 2015. DOI: 10.1007/978-1-4899-7637-6_10. URL: https://link.springer.com/chapter/10.1007/978-1-4899-7637-6_10.

[101] Thi Ngoc Trang Tran et al. “Do Users Appreciate Explanations of Recommendations? An Analysis in the Movie Domain”. In: ACM, Sept. 2021, pp. 645–650. ISBN: 9781450384582. DOI: 10.1145/3460231.3478859.

[102] James Turland et al. “Nudging Towards security: Developing an Application for Wireless Network Selection for Android Phones”. In: *Proceedings of the 2015 British HCI Conference on - British HCI ’15* (2015), pp. 193–201. DOI: 10.1145/2783446.2783588. URL: http://dl.acm.org/citation.cfm?id=2783446.2783588{\%}5Cnhttp://dl.acm.org/citation.cfm?doid=2783446.2783588.

[103] Blase Ur et al. “How does your password measure up? The effect of strength meters on password creation”. In: *Security’12 Proceedings of the 21st USENIX conference on Security symposium* (2012), pp. 5–5. URL: https://www.usenix.org/system/files/conference/usenixsecurity12/sec12-final209.pdf.

[104] C. Utz et al. “(Un)informed Consent: Studying GDPR consent notices in the field”. In: 2019, pp. 973–990. ISBN: 9781450367479. DOI: 10.1145/3319535.3354212.

[105] Paul Veyne and Arthur Goldhammer. *From pagan Rome to Byzantium. Vol. 1 of A history of private life.* 1987.

[106] Ivo Vlaev. "How different are real and hypothetical decisions? Overestimation, contrast and assimilation in social interaction". In: *Journal of Economic Psychology* 33 (5 Oct. 2012), pp. 963–972. ISSN: 01674870. DOI: `10.1016/j.joep.2012.05.005`.

[107] Yang Wang et al. "A field trial of privacy nudges for facebook". In: *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14* (2014), pp. 2367–2376. DOI: `10.1145/2556288.2557413`. URL: `http://dl.acm.org/citation.cfm?id=2556288.2557413`.

[108] Taiyun Wei and Viliam Simko. *R package "corrplot": Visualization of a Correlation Matrix.* 2017. URL: `https://github.com/taiyun/corrplot`.

[109] Alan Westin and Oscar Ruebhausen. *Privacy and freedom.* 1967.

[110] Hadley Wickham and Romain Francois. *dplyr: A Grammar of Data Manipulation.* 2016. URL: `https://cran.r-project.org/package=dplyr`.

[111] Jierui Xie, Bart Piet Knijnenburg, and Hongxia Jin. "Location sharing privacy preference". In: ACM, Feb. 2014. ISBN: 9781450321846. DOI: `10.1145/2557500.2557504`. URL: `https://dl.acm.org/doi/10.1145/2557500.2557504`.

[112] Hengshu Zhu et al. "Mobile app recommendations with security and privacy awareness". In: ACM, Aug. 2014. ISBN: 9781450329569. DOI: `10.1145/2623330.2623705`. URL: `https://dl.acm.org/doi/10.1145/2623330.2623705`.

Appendix A

Appendices for Tailoring Privacy Nudges to Individual Differences

A.1 Text of Experimental Manipulations

A.1.1 Study 1: Auto Checkout Disclosure

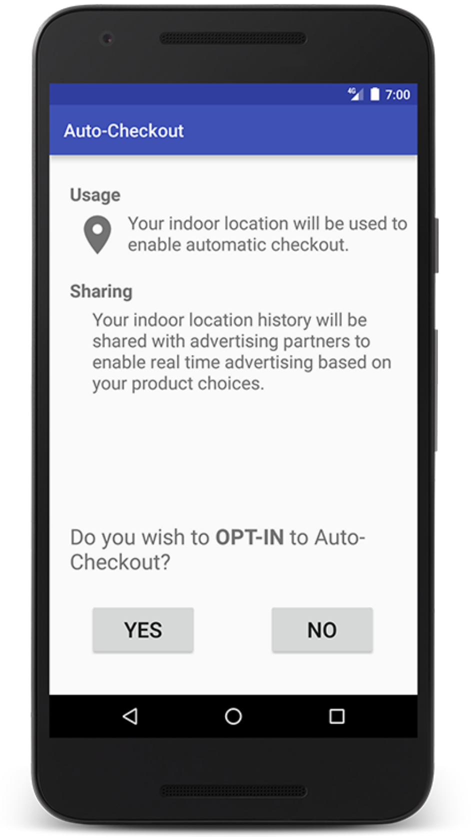
'Opt-In' Condition

Imagine a new supermarket in your town has implemented sensors which allow it to determine your location within the store based on the position of your smartphone. This technology is used by the supermarket to enable a new service called "Auto-Checkout", which lets you skip checkout lines by using your location to identify which products you intend to purchase and charging your credit card. The supermarket may also use your location data to provide you with ads for products based on your previous purchases. These functionalities are implemented through a smartphone app that you have installed on your phone. The "Auto-Checkout" system has been tested by the supermarket to ensure accuracy, reliability, and security. As you enter the supermarket, the following message is displayed to you on your smartphone screen.

See Figure A.1

'Opt-Out' Condition

Imagine a new supermarket in your town has implemented sensors which allow it to determine your location within the store based on the position of your smartphone. This technology is used by the supermarket to enable a new service called "Auto-Checkout", which lets you skip checkout lines by using your location to identify which products you intend to purchase and charging your credit card. The supermarket may also use your location data to provide you with ads for products based on your previous purchases. These functionalities are implemented through a smartphone app that you have installed on your phone. The "Auto-Checkout" system has been tested



7:00

Auto-Checkout

Usage

Your indoor location will be used to enable automatic checkout.

Sharing

Your indoor location history will be shared with advertising partners to enable real time advertising based on your product choices.

Do you wish to **OPT-IN** to Auto-Checkout?

Figure A.1: Interactive phone screen for 'Opt-In' Condition in Study 1

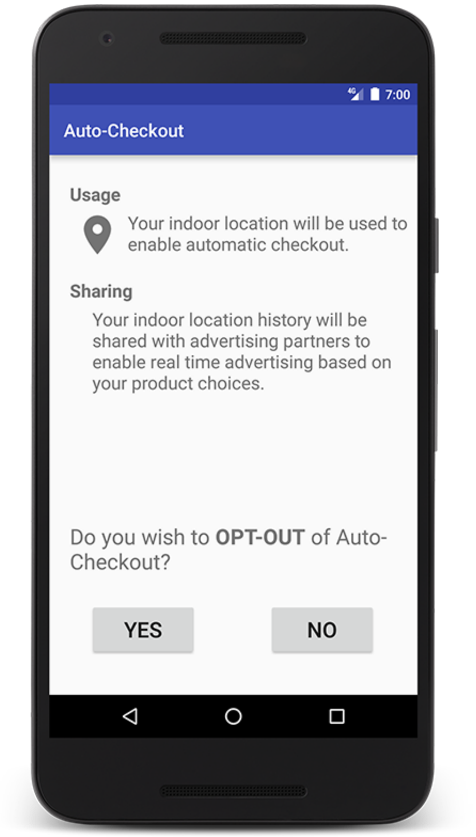
by the supermarket to ensure accuracy, reliability, and security. As you enter the supermarket, the following message is displayed to you on your smartphone screen. *See Figure A.2*

A.1.2 Study 2 & 3: Ethical Behavior Questions Disclosure

Framing Nudge Allow Condition

Before we ask you the ethical behavior questions, we want to determine your preferences over the sharing of your responses. **You can have your responses to the ethical behavior questions, along with your Mechanical Turk ID, shared with researchers outside of our study team. If you consent to sharing, we will make this data available to researchers outside of our study team.** These researchers will use your responses as part of future published research studies. We show your Mechanical Turk ID as Participant's Mechanical Turk ID.

Please note that your Mechanical Turk ID may not be anonymous. Recent information has shown that your Mechanical Turk ID may be linked to information that can



7:00

Auto-Checkout

Usage

Your indoor location will be used to enable automatic checkout.

Sharing

Your indoor location history will be shared with advertising partners to enable real time advertising based on your product choices.

Do you wish to **OPT-OUT** of Auto-Checkout?

Figure A.2: Interactive phone screen for 'Opt-Out' Condition in Study 1

identify you such as your full name and Amazon purchase history. While we do not access or use this information in our study, other researchers may use this information in future published research studies.

Allow your responses to the ethical behavior questions to be shared with **researchers outside of our study team, along with your Mechanical Turk ID?**

Framing Nudge Prohibit Condition

Before we ask you the ethical behavior questions, we want to determine your preferences over the sharing of your responses. **You can have your responses to the ethical behavior questions, along with your Mechanical Turk ID, shared with researchers outside of our study team. If you consent to sharing, we will make this data available to researchers outside of our study team.** These researchers will use your responses as part of future published research studies. We show your Mechanical Turk ID as Participant's Mechanical Turk ID.

Please note that your Mechanical Turk ID may not be anonymous. Recent information has shown that your Mechanical Turk ID may be linked to information that can identify you such as your full name and Amazon purchase history. While we do not access or use this information in our study, other researchers may use this information in future published research studies.

Prohibit your responses to the ethical behavior questions from being shared with **researchers outside of our study team, along with your Mechanical Turk ID?**

Social Norms Nudge 'High Norms' Condition

Before we ask you the ethical behavior questions, we want to determine your preferences over the sharing of your responses. **You can have your responses to the ethical behavior questions, along with your Mechanical Turk ID, shared with researchers outside of our study team. If you consent to sharing, we will make this data available to researchers outside of our study team.** These researchers will use your responses as part of future published research studies. We show your Mechanical Turk ID as Participant's Mechanical Turk ID.

Please note that your Mechanical Turk ID may not be anonymous. Recent information has shown that your Mechanical Turk ID may be linked to information that can identify you such as your full name and Amazon purchase history. While we do not access or use this information in our study, other researchers may use this information in future published research studies. **In our past studies, 73% of participants chose to allow their responses to be shared with researchers outside of our study team.**

Please select a sharing preference below:

- **Allow** my responses and Mechanical Turk ID to be shared with researchers outside of the study team
- **Prohibit** my responses and Mechanical Turk ID from being shared with researchers outside of the study team

Social Norms Nudge 'Low Norms' Condition

Before we ask you the ethical behavior questions, we want to determine your preferences over the sharing of your responses. **You can have your responses to the ethical behavior questions, along with your Mechanical Turk ID, shared with researchers outside of our study team. If you consent to sharing, we will make this data available to researchers outside of our study team.** These researchers will use your responses as part of future published research studies. We show your Mechanical Turk ID as Participant's Mechanical Turk ID.

Please note that your Mechanical Turk ID may not be anonymous. Recent information has shown that your Mechanical Turk ID may be linked to information that can identify you such as your full name and Amazon purchase history. While we do not access or use this information in our study, other researchers may use this information in future published research studies. **In our past studies, 31% of participants chose to allow their responses to be shared with researchers outside of our study team.**

Please select a sharing preference below:

- **Allow** my responses and Mechanical Turk ID to be shared with researchers outside of the study team
- **Prohibit** my responses and Mechanical Turk ID from being shared with researchers outside of the study team

A.2 Psychometric Variable Scales

A.2.1 Resistance to Framing

Questions are from the Adult Decision-Making Competence inventory. The complete set of questions used in our studies can be found at: `http://www.sjdm.org/dmidi/Adult_-_Decision_Making_Competence.html`.

Part 1 Sample Questions

Each of the following problems presents a choice between two options. Each problem is presented with a scale ranging from 1 (representing one option) through 6 (representing the other option). For each item, please select the number on the scale that best reflects your relative preference between the two options.

Q1: Imagine that recent evidence has shown that a pesticide is threatening the lives of 1,200 endangered animals. Two response options have been suggested:

- If Option A is used, 600 animals will be saved for sure.
- If Option B is used, there is a 75% chance that 800 animals will be saved, and a 25% chance that no animals will be saved.

Which option do you recommend to use?

Q2: Because of changes in tax laws, you may get back as much as $1200 in income tax. Your accountant has been exploring alternative ways to take advantage of this situation. He has developed two plans:

- If Plan A is adopted, you will get back $400 of the possible $1200.
- If Plan B is adopted, you have a 33% chance of getting back all $1200, and a 67% chance of getting back no money.

Which plan would you use?

Part 2 Sample Questions

Each of the following problems presents a choice between two options. Each problem is presented with a scale ranging from 1 (representing one option) through 6 (representing the other option). For each item, please select the number on the scale that best reflects your relative preference between the two options.

Q1: Imagine that recent evidence has shown that a pesticide is threatening the lives of 1,200 endangered animals. Two response options have been suggested:

- If Option A is used, 600 animals will be lost for sure.
- If Option B is used, there is a 75% chance that 400 animals will be lost, and a 25% chance that 1,200 animals will be lost.

Which option do you recommend to use?

Q2: Because of changes in tax laws, you may get back as much as $1200 in income tax. Your accountant has been exploring alternative ways to take advantage of this situation. He has developed two plans:

- If Plan A is adopted, you will lose $800 of the possible $1200.
- If Plan B is adopted, you have a 33% chance of losing none of the money, and a 67% chance of losing all $1200.

Which plan would you use?

A.2.2 Recognizing Social Norms

Questions are from the Adult Decision-Making Competence inventory. The complete set of questions used in our studies can be found at: `http://www.sjdm.org/dmidi/Adult_-_Decision_Making_Competence.html`.

Part 1 Sample Questions

The following problems ask whether it is sometimes OK to do different things. For each question, please indicate whether *in your opinion* the answer is yes or no.

Do you think it is sometimes OK...

- to steal under certain circumstances?
- to smoke cigarettes?
- to commit a crime which could put you in jail?
- to keep things you find in the street?
- to experiment with marijuana?

Part 2 Sample Questions

The following problems ask out of *100 people your age*, how many would say that it is sometimes OK to do different things. For each question, please select a number between 0 (meaning *no one* thinks that it is sometimes OK) and 100 (meaning *everyone* thinks that it is sometimes OK).

Out of 100 people your age, how many would say it is sometimes OK ...

- to steal under certain circumstances?
- to smoke cigarettes?
- to commit a crime which could put you in jail?
- to keep things you find in the street?
- to experiment with marijuana?

A.3 Post-Hoc Power Analysis Results

A.3.1 Social Norms Nudge and Recognizing Social Norms Score

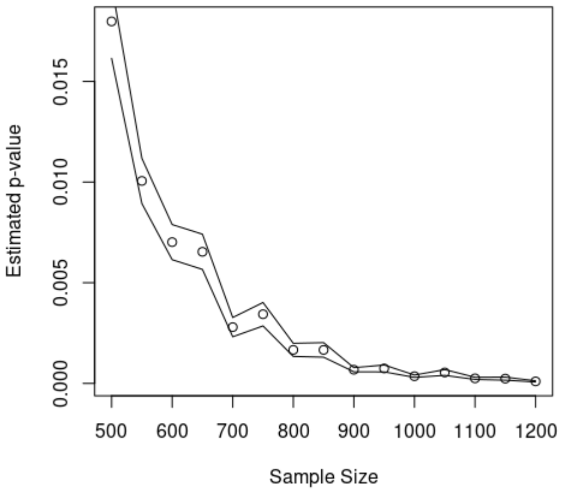


Figure A.3: Estimated p-values for different sample sizes for the Social Norms Nudge and Recognizing Social Norms score with 95% confidence interval

A.3.2 Framing Nudge and Big 5 Extraversion Score

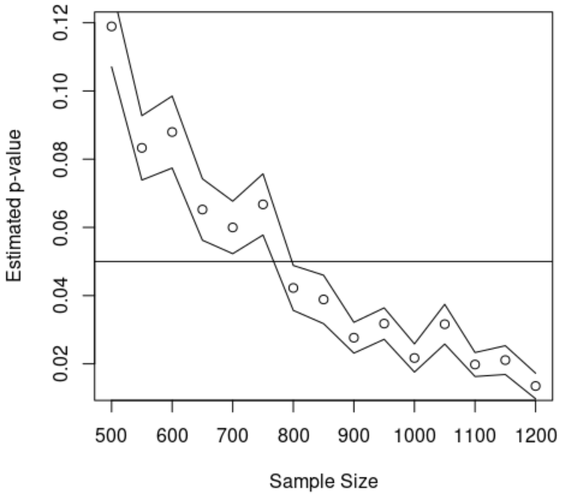


Figure A.4: Estimated p-values for different sample sizes for the Framing Nudge and Big 5 Extraversion score with 95% confidence interval

A.3.3 Framing Nudge and Big 5 Conscientiousness Score

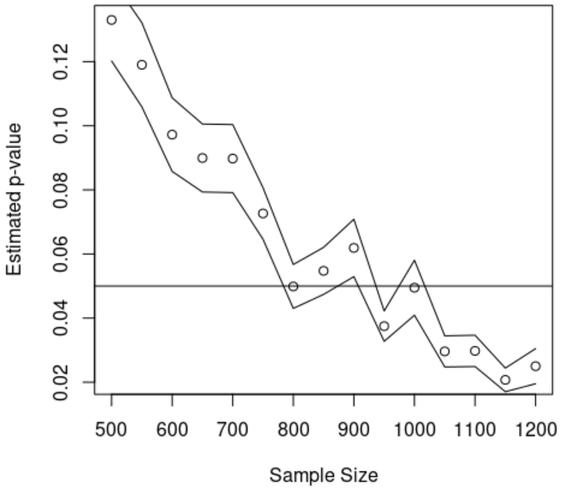


Figure A.5: Estimated p-values for different sample sizes for the Framing Nudge and Big 5 Extraversion score with 95% confidence interval

A.4 Additional Regression Tables

A.4.1 Study 1 Regression Tables

Table A.1: Framing Score and Log-ADR Score Logit Regressions for Study 1

	Dependent variable:							
	disclosure							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Allow Frame	3.780	3.959	3.884	3.909	0.639	0.632	0.683	0.745
	(3.396)	(3.436)	(3.447)	(3.496)	(0.544)	(0.544)	(0.554)	(0.560)
Framing Score	0.382	0.358	0.352	0.266				
	(0.592)	(0.594)	(0.598)	(0.607)				
CFIP Score		−0.179	−0.220	−0.238		−0.097	−0.130	−0.143
		(0.181)	(0.188)	(0.192)		(0.184)	(0.189)	(0.190)
Age			0.009	0.007			0.008	0.006
			(0.018)	(0.019)			(0.019)	(0.019)
Female			0.414	0.391			0.506	0.493
			(0.374)	(0.378)			(0.381)	(0.383)
Years of Education				0.141				0.076
				(0.093)				(0.096)
Allow Frame*Framing Score	−0.761	−0.814	−0.773	−0.768				
	(0.871)	(0.882)	(0.885)	(0.898)				
Log-ADR Score					1.064**	1.024*	1.112**	1.001*
					(0.526)	(0.525)	(0.550)	(0.565)
Allow Frame*Log-ADR Score					−0.232	−0.221	−0.311	−0.250
					(0.763)	(0.760)	(0.783)	(0.787)
Constant	−1.703	−0.540	−0.797	−2.352	0.408	0.967	0.682	−0.381
	(2.306)	(2.589)	(2.612)	(2.835)	(0.380)	(1.123)	(1.229)	(1.818)
Observations	143	143	143	143	143	143	143	143
Log Likelihood	−95.133	−94.630	−93.792	−92.626	−91.910	−91.768	−90.703	−90.389
Akaike Inf. Crit.	198.265	199.260	201.585	201.252	191.820	193.536	195.406	196.778
Note:								*p<0.1; **p<0.05; ***p<0.01

A.4.2 Study 2 Regression Tables

Table A.2: Social Norms Nudge and Recognizing Social Norms Regressions for Study 2

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Recognizing Social Norms	−1.744***	−1.318**	−1.194**	−1.232**
	(0.538)	(0.564)	(0.573)	(0.574)
High Norms Condition	−0.429	−0.281	−0.293	−0.322
	(0.395)	(0.412)	(0.419)	(0.420)
IUIPC		−0.591***	−0.536***	−0.527***
		(0.125)	(0.127)	(0.128)
Age			−0.020**	−0.021**
			(0.010)	(0.010)
Female			−0.086	−0.083
			(0.207)	(0.208)
African American			−0.023	−0.012
			(0.360)	(0.363)
Hispanic			0.205	0.222
			(0.404)	(0.405)
Asian			0.491	0.489
			(0.489)	(0.493)
Other Race			0.197	0.147
			(0.620)	(0.621)
High School				−12.697
				(535.411)
Associate Degree				−12.650
				(535.411)
Bachelor's Degree				−12.768
				(535.411)
Advanced Degree				−13.118
				(535.411)
Recognizing Social Norms*High Norms Condition	2.542***	2.294***	2.400***	2.432***
	(0.743)	(0.772)	(0.783)	(0.784)
Constant	0.572**	3.870***	4.225***	16.968
	(0.283)	(0.764)	(0.808)	(535.411)
Observations	453	453	453	453
Log Likelihood	−298.305	−285.683	−282.242	−281.000
Akaike Inf. Crit.	604.611	581.366	586.485	591.999
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table A.3: Framing Nudge and Big 5 Extraversion Regressions for Study 2

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Big 5 Extraversion	−0.022	−0.045	−0.058	−0.056
	(0.133)	(0.137)	(0.139)	(0.140)
Allow Condition	−0.260	−0.336	−0.400	−0.398
	(0.595)	(0.605)	(0.612)	(0.616)
IUIPC		-0.497^{***}	-0.480^{***}	-0.485^{***}
		(0.118)	(0.121)	(0.122)
Age			−0.009	−0.008
			(0.009)	(0.009)
Female			0.205	0.200
			(0.202)	(0.203)
African American			0.170	0.162
			(0.367)	(0.369)
Hispanic			0.056	0.027
			(0.360)	(0.365)
Asian			0.417	0.451
			(0.464)	(0.472)
Other Race			0.136	0.130
			(0.650)	(0.648)
Associate Degree				0.166
				(0.327)
Bachelor's Degree				0.274
				(0.302)
Advanced Degree				−0.086
				(0.391)
Other Education				13.218
				(535.411)
Big 5 Extraversion*Allow Condition	0.433^{**}	0.456^{**}	0.486^{**}	0.487^{**}
	(0.197)	(0.200)	(0.203)	(0.204)
Constant	−0.336	2.657^{***}	2.796^{***}	2.615^{***}
	(0.404)	(0.824)	(0.846)	(0.864)
Observations	482	482	482	482
Log Likelihood	−314.760	−304.871	−303.238	−301.875
Akaike Inf. Crit.	637.520	619.742	628.477	633.749
Note:		*p<0.1; **p<0.05; ***p<0.01		

Table A.4: Framing Nudge and Big 5 Conscientiousness Regressions for Study 2

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Big 5 Conscientiousness	0.137	0.280	0.282	0.258
	(0.177)	(0.188)	(0.192)	(0.193)
Allow Condition	3.041***	2.992***	2.915***	2.849***
	(1.023)	(1.053)	(1.058)	(1.061)
IUIPC		−0.494***	−0.484***	−0.485***
		(0.120)	(0.122)	(0.123)
Age			−0.008	−0.007
			(0.009)	(0.009)
Female			0.159	0.157
			(0.200)	(0.201)
African American			0.185	0.171
			(0.366)	(0.368)
Hispanic			−0.022	−0.052
			(0.361)	(0.366)
Asian			0.343	0.355
			(0.460)	(0.469)
Other Race			0.114	0.094
			(0.649)	(0.647)
Associate Degree				0.116
				(0.327)
Bachelor's Degree				0.283
				(0.302)
Advanced Degree				−0.013
				(0.387)
Other Education				13.081
				(535.411)
Big 5 Conscientiousness*Allow Condition	−0.514**	−0.506*	−0.482*	−0.464*
	(0.253)	(0.260)	(0.262)	(0.262)
Constant	−0.933	1.420	1.530	1.425
	(0.706)	(0.930)	(0.949)	(0.963)
Observations	482	482	482	482
Log Likelihood	−316.376	−307.109	−305.940	−304.746
Akaike Inf. Crit.	640.752	624.218	633.881	639.492
Note:				*p<0.1; **p<0.05; ***p<0.01

A.4.3 Study 3 Regression Tables

Table A.5: Social Norms Nudge and Recognizing Social Norms Regressions for Study 3

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Recognizing Social Norms	−1.233***	−0.985***	−1.031***	−1.054***
	(0.338)	(0.364)	(0.374)	(0.377)
High Norms Condition	0.438	0.309	0.374	0.363
	(0.269)	(0.291)	(0.299)	(0.300)
IUIPC		−0.924***	−0.952***	−0.953***
		(0.097)	(0.100)	(0.101)
Age			0.012*	0.011*
			(0.006)	(0.006)
Female			−0.291**	−0.303**
			(0.147)	(0.148)
Non-Binary			0.911	0.859
			(1.252)	(1.250)
African American			−0.408	−0.414
			(0.263)	(0.265)
Hispanic			0.110	0.096
			(0.320)	(0.320)
Asian			−0.837***	−0.825**
			(0.324)	(0.325)
Other Race			0.168	0.144
			(0.488)	(0.490)
High School				0.590
				(1.038)
Associate Degree				0.513
				(1.020)
Bachelor's Degree				0.374
				(1.016)
Advanced Degree				0.518
				(1.029)
Recognizing Social Norms*High Norms Condition	0.529	0.916*	0.800	0.818
	(0.481)	(0.520)	(0.532)	(0.534)
Constant	0.501***	5.829***	5.781***	5.367***
	(0.189)	(0.611)	(0.642)	(1.160)
Observations	936	936	936	936
Log Likelihood	−618.414	−560.791	−551.897	−551.202
Akaike Inf. Crit.	1,244.828	1,131.581	1,127.794	1,134.404
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.6: Framing Nudge and Big 5 Extraversion Regressions for Study 3

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Big 5 Extraversion	−0.096	−0.125	−0.099	−0.062
	(0.110)	(0.115)	(0.116)	(0.117)
Allow Condition	0.195	0.256	0.246	0.283
	(0.452)	(0.468)	(0.471)	(0.475)
IUIPC		-0.508^{***}	-0.483^{***}	-0.500^{***}
		(0.069)	(0.070)	(0.071)
Age			-0.010^{*}	−0.010
			(0.006)	(0.006)
Female			−0.007	−0.016
			(0.139)	(0.140)
Non-Binary			0.984	1.206
			(1.230)	(1.223)
African American			−0.295	−0.350
			(0.235)	(0.237)
Hispanic			−0.094	−0.107
			(0.280)	(0.285)
Asian			−0.169	−0.088
			(0.266)	(0.271)
Other Race			−0.368	−0.476
			(0.667)	(0.669)
Associate Degree				1.509
				(0.985)
Bachelor's Degree				1.826^{*}
				(0.974)
Advanced Degree				1.354
				(0.969)
Other Education				1.063
				(0.979)
Big 5 Extraversion*Allow Condition	0.302^{**}	0.281^{*}	0.289^{*}	0.287^{*}
	(0.151)	(0.157)	(0.158)	(0.159)
Constant	−0.421	2.580^{***}	2.796^{***}	1.331
	(0.329)	(0.532)	(0.554)	(1.093)
Observations	992	992	992	992
Log Likelihood	−648.681	−618.239	−615.658	−607.651
Akaike Inf. Crit.	1,305.363	1,246.478	1,255.315	1,247.302
Note:				*p<0.1; **p<0.05; ***p<0.01

Table A.7: Framing Nudge and Big 5 Conscientiousness Regressions for Study 3

	Dependent variable:			
	disc			
	(1)	(2)	(3)	(4)
Big 5 Conscientiousness	-0.380^{***}	-0.198	-0.176	-0.141
	(0.121)	(0.127)	(0.128)	(0.130)
Allow Condition	-0.014	0.238	0.230	0.364
	(0.659)	(0.685)	(0.686)	(0.692)
IUIPC		-0.491^{***}	-0.473^{***}	-0.496^{***}
		(0.070)	(0.071)	(0.072)
Age			-0.009	-0.009
			(0.006)	(0.006)
Female			-0.005	-0.015
			(0.139)	(0.140)
Non-Binary			1.078	1.273
			(1.214)	(1.202)
African American			-0.235	-0.279
			(0.231)	(0.233)
Hispanic			-0.080	-0.091
			(0.280)	(0.284)
Asian			-0.162	-0.084
			(0.266)	(0.271)
Other Race			-0.370	-0.472
			(0.669)	(0.671)
Associate Degree				1.396
				(0.982)
Bachelor's Degree				1.707^{*}
				(0.971)
Advanced Degree				1.267
				(0.966)
Other Education				0.967
				(0.976)
Big 5 Conscientiousness*Allow Condition	0.284^{*}	0.216	0.221	0.194
	(0.169)	(0.176)	(0.176)	(0.177)
Constant	0.747	2.877^{***}	3.065^{***}	1.702
	(0.466)	(0.575)	(0.591)	(1.118)
Observations	992	992	992	992
Log Likelihood	-645.753	-618.697	-616.582	-609.227
Akaike Inf. Crit.	1,299.507	1,247.394	1,257.163	1,250.455
Note:		*p<0.1; **p<0.05; ***p<0.01		

Appendix B

Appendices for Explanations for Mobile Privacy Recommendations

B.1 Full Regression Tables

Table B.1: Allow Recommendations Model 1

	Dependent variable:				
	intent				
	(1)	(2)	(3)	(4)	(5)
Social Explanation	0.022	−0.363	−0.474	−0.471	−0.337
	(0.290)	(0.378)	(0.396)	(0.397)	(0.421)
Behavioral Explanation	0.333	0.294	0.205	0.206	0.294
	(0.290)	(0.380)	(0.399)	(0.399)	(0.434)
Hybrid Explanation	0.153	0.035	−0.072	−0.064	0.185
	(0.300)	(0.385)	(0.399)	(0.400)	(0.429)
Rec Quality		1.440***	1.239***	1.238***	1.226***
		(0.143)	(0.147)	(0.147)	(0.157)
Trust			0.821***	0.823***	0.878***
			(0.170)	(0.170)	(0.183)
Video Conf Apps Installed				−0.166	0.129
				(0.592)	(0.642)
IUIPC					−0.685***
					(0.194)
Age					0.023*
					(0.013)
Male					0.668**
					(0.326)
Doctorate Degree					−13.012
					(882.744)
Graduate Degree					−12.661
					(882.744)
High School Diploma					−12.300
					(882.743)
Secondary Education (GED)					−13.889
					(882.744)
Technical/Community College					−13.020
					(882.744)
Undergraduate Degree					−12.400
					(882.743)
Constant	−0.105	−4.090***	−6.233***	−6.085***	8.836
	(0.205)	(0.477)	(0.720)	(0.889)	(882.745)
Observations	372	372	372	372	372
Log Likelihood	−256.987	−172.125	−159.555	−159.515	−146.382
Akaike Inf. Crit.	521.974	354.251	331.109	333.030	324.764
Note:					*p<0.1; **p<0.05; ***p<0.01

Table B.2: Allow Recommendations Model 2

	Dependent variable:					
	intent					
	(1)	(2)	(3)	(4)	(5)	(6)
Social Explanation	−0.131	−0.382	−0.334	−0.364	−0.367	−0.479
	(0.299)	(0.379)	(0.391)	(0.398)	(0.398)	(0.417)
Behavioral Explanation	0.180	0.253	0.275	0.287	0.283	0.141
	(0.299)	(0.382)	(0.389)	(0.402)	(0.404)	(0.426)
Rec Quality		1.389***	1.275***	1.166***	1.166***	1.125***
		(0.161)	(0.164)	(0.166)	(0.166)	(0.176)
Exp Quality			0.512***	0.234	0.233	0.425**
			(0.148)	(0.174)	(0.174)	(0.192)
Trust				0.770***	0.771***	0.713***
				(0.222)	(0.222)	(0.232)
Video Conf Apps Installed					−0.095	0.221
					(0.722)	(0.787)
IUIPC						−0.624***
						(0.230)
Age						0.015
						(0.015)
Male						0.911**
						(0.378)
Doctorate Degree						−13.598
						(882.744)
Graduate Degree						−13.583
						(882.744)
High School Diploma						−13.117
						(882.744)
Secondary Education (GED)						−13.861
						(882.744)
Technical/Community College						−13.680
						(882.744)
Undergraduate Degree						−13.347
						(882.744)
Constant	0.048	−3.908***	−5.202***	−6.665***	−6.575***	9.043
	(0.218)	(0.538)	(0.704)	(0.888)	(1.117)	(882.746)
Observations	277	277	277	277	277	277
Log Likelihood	−191.270	−131.385	−125.163	−118.654	−118.645	−110.528
Akaike Inf. Crit.	388.539	270.771	260.326	249.307	251.290	253.056
Note:						*p<0.1; **p<0.05; ***p<0.01

Table B.3: Deny Recommendations Model 1

	Dependent variable:				
	intent				
	(1)	(2)	(3)	(4)	(5)
Social Explanation	-0.435	-0.364	-0.377	-0.394	-0.355
	(0.657)	(0.667)	(0.668)	(0.669)	(0.691)
Behavioral Explanation	-0.946	-1.074^{*}	-1.077^{*}	-1.105^{*}	-1.022
	(0.604)	(0.615)	(0.616)	(0.617)	(0.630)
Hybrid Explanation	-0.916	-1.227^{**}	-1.242^{**}	-1.280^{**}	-1.111^{*}
	(0.604)	(0.623)	(0.625)	(0.628)	(0.647)
Rec Quality		0.551^{***}	0.544^{***}	0.541^{***}	0.524^{***}
		(0.140)	(0.142)	(0.143)	(0.153)
Trust			0.072	0.090	-0.046
			(0.219)	(0.221)	(0.243)
Video Conf Apps Installed				-0.577	-0.517
				(0.632)	(0.668)
IUIPC					0.464^{**}
					(0.207)
Age					0.010
					(0.018)
Male					-0.386
					(0.420)
Doctorate Degree					15.644
					(910.972)
Graduate Degree					1.128
					(1.281)
High School Diploma					1.594
					(1.225)
Secondary Education (GED)					0.532
					(1.602)
Technical/Community College					1.201
					(1.238)
Undergraduate Degree					1.660
					(1.228)
Constant	3.534^{***}	1.451^{**}	1.203	1.669	-2.316
	(0.507)	(0.691)	(1.021)	(1.152)	(2.049)
Observations	570	570	570	570	570
Log Likelihood	-115.692	-108.803	-108.750	-108.275	-103.532
Akaike Inf. Crit.	239.385	227.607	229.501	230.549	239.065
Note:			*p<0.1; **p<0.05; ***p<0.01		

Table B.4: Deny Recommendations Model 2

	Dependent variable:					
	intent					
	(1)	(2)	(3)	(4)	(5)	(6)
Social Explanation	0.481	0.888	0.854	0.845	0.890	0.654
	(0.530)	(0.565)	(0.568)	(0.568)	(0.571)	(0.611)
Behavioral Explanation	−0.030	0.166	0.134	0.139	0.162	0.028
	(0.463)	(0.480)	(0.483)	(0.483)	(0.486)	(0.510)
Rec Quality		0.574***	0.579***	0.573***	0.578***	0.547***
		(0.153)	(0.153)	(0.155)	(0.156)	(0.168)
Exp Quality			−0.126	−0.158	−0.142	−0.108
			(0.202)	(0.232)	(0.233)	(0.258)
Trust				0.084	0.098	−0.077
				(0.284)	(0.284)	(0.316)
Video Conf Apps Installed					−0.914	−1.071
					(0.762)	(0.814)
IUIPC						0.435*
						(0.223)
Age						0.003
						(0.019)
Male						−0.365
						(0.460)
Doctorate Degree						16.149
						(964.921)
Graduate Degree						1.845
						(1.396)
High School Diploma						1.852
						(1.286)
Secondary Education (GED)						0.342
						(1.707)
Technical/Community College						1.603
						(1.324)
Undergraduate Degree						2.517*
						(1.325)
Constant	2.617***	0.126	0.570	0.374	1.026	−2.534
	(0.328)	(0.715)	(1.017)	(1.206)	(1.357)	(2.148)
Observations	429	429	429	429	429	429
Log Likelihood	−97.500	−91.170	−90.973	−90.929	−90.033	−84.711
Akaike Inf. Crit.	200.999	190.341	191.945	193.859	194.066	201.421
Note:						*p<0.1; **p<0.05; ***p<0.01

B.2 Privacy Questions

Imagine that you have installed a Privacy Assistant app on your smartphone. This app asks you questions about your privacy preferences, and then, based on your answers, provides recommendations for permission settings (whether different types of personal data can be accessed) for apps on your smartphone.

On the next screen, you will be asked to answer questions from the Privacy Assistant app about your privacy preferences. Please answer them. After examining your responses, we will provide you with a permission setting recommendation generated by the Privacy Assistant app.

Q1: In general, how comfortable or uncomfortable do you feel with **Communication** apps (such as Google Chrome, Gmail, or WhatsApp) accessing your **Calendar**?

Q2: In general, how comfortable or uncomfortable do you feel with **Entertainment** apps (such as Google Play Games, Netflix, or Mi Video) accessing your **Contacts**?

Q3: In general, how comfortable or uncomfortable do you feel with **Productivity** apps (such as Google Drive, Google Calendar, or Microsoft OneDrive) accessing your **Camera**?

Q4: In general, how comfortable or uncomfortable do you feel with **Social** apps (such as Facebook, Instagram, or Twitter) accessing your **Location**?

Q5: In general, how comfortable or uncomfortable do you feel with **Music and Audio** apps (such as YouTube Music, Spotify, or Samsung Music) accessing your **Microphone**?

Response options: Very uncomfortable, Moderately uncomfortable, Somewhat uncomfortable, Somewhat comfortable, Moderately comfortable, Very comfortable

B.3 Followup Questions Text

B.3.1 Memory Recall Questions

To conclude our study, please answer the following additional questions.

Please recall the mobile app permission recommendation and the recommended permission selection it provided to the best of your ability.

Q1: What was the name of the mobile app contained in the recommendation?
Response options: VideoMeet, MailCenter, VideoWatch, RideFinder, AccuMaps, TrailMaps

Q2: Which permission did the recommendation provide a permission selection for?
Response options: Phone, SMS, Calendar, Location, Camera, Microphone

Q3: What permission selection did the recommendation suggest making?
Response options: Allow the app access to the permission, Deny the app access to the permission, No permission selection was recommended

B.3.2 Recommendation Quality Questions

Q1 (Allow Recommendation): Consider all aspects of your experience using the personalized privacy assistant. What factors were most important in your decision of whether or not to allow the VideoMeet app access to your Location?

Q1 (Deny Recommendation): Consider all aspects of your experience using the personalized privacy assistant. What factors were most important in your decision of whether or not to deny the VideoMeet app access to your Location?

Q2 (Allow Recommendation): Please indicate how strongly you agree or disagree with the statements below regarding the above recommended permission selection. When answering these questions, please respond according to how you felt when you viewed the recommended permission selection in the privacy assistant app: By allowing the VideoMeet app access to my Location, I thought that the types of data accessible by the VideoMeet app would better align with my privacy preferences. *Response options: Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree*

Q2 (Deny Recommendation): Please indicate how strongly you agree or disagree with the statements below regarding the above recommended permission selection. When answering these questions, please respond according to how you felt when you viewed the recommended permission selection in the privacy assistant app: By denying the VideoMeet app access to my Location, I thought that the types of data accessible by the VideoMeet app would better align with my privacy preferences. *Response options: Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree*

Q3: Please briefly explain how the recommendation agrees or conflicts with your privacy preferences.

Q4: How necessary do you think it is for the VideoMeet app to have access to your location while in use? *Response options: Extremely necessary, Very necessary, Moderately necessary, Slightly necessary, Not at all necessary, I don't know*

B.3.3 Explanation Quality Questions

Q1: How important was the explanation when deciding whether or not you wanted to adopt the change to your permission settings recommended by the privacy assistant? *Response options: Extremely important, Very important, Moderately important, Slightly important, Not at all important*

Q2: Please briefly explain how the explanation impacted your decision whether or not to [allow/deny] the VideoMeet app access to your Location.

Q3: Please indicate how strongly you agree or disagree with the statements below regarding the above explanation:
— The explanation helped me understand why the privacy assistant made the particular recommendation I received.
— The explanation helped me to achieve a more satisfactory outcome when deciding whether or not to adopt the recommendation I received.
— The explanation helped me to decide whether or not to adopt the recommendation I received more quickly.

— Understanding how the personalized privacy assistant generates its recommendations was important to me when deciding whether or not to adopt the recommendation I received.
Response options: Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree

B.3.4 Trust in Privacy Assistant Questions

Q1: Please indicate how strongly you agree or disagree with the statements below regarding the privacy assistant app:
— I believe that the privacy assistant is honest about the recommendations it provides.
— I believe that the privacy assistant aims to make recommendations that are in my best interests.
— I believe that the privacy assistant is capable of making recommendations that align with my privacy preferences.
Response options: Strongly agree, Somewhat agree, Neither agree nor disagree, Somewhat disagree, Strongly disagree

B.3.5 Video Conferencing App Preferences Questions

Q1: What operating system (OS) does your mobile device run? *Response options: Android, iOS, Other (open-response), I don't know*

Q2: Which of the following video conferencing apps do you have installed on your mobile device? Please select all that apply. *Response options: Zoom, Skype, FaceTime, Microsoft Teams, Google Meet, Google Duo, Cisco Webex, Other (open-response), I don't have any video conferencing apps installed on my mobile device*

Q3: How often do you use video conferencing apps on your mobile device? *Response options: Never, Several times a year, Several times a month, Several times a week, Daily, I don't have any video conferencing apps installed on my mobile device*

Q4: How important are the video conferencing apps on your mobile device to completing your day-to-day responsibilities? *Response options: Extremely important, Very important, Moderately important, Slightly important, Not at all important, I don't have any video conferencing apps installed on my mobile device*

Q5: On average, how necessary do you think the Location permission is to the function of the video conferencing app(s) on your mobile device? *Response options: Extremely necessary, Very necessary, Moderately necessary, Slightly necessary, Not at all necessary, I don't have any video conferencing apps installed on my mobile device*