Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs

Jessica Colnago, *Google;* Lorrie Faith Cranor and Alessandro Acquisti, *Carnegie Mellon University;* Kate Hazel Stanton, *University of Pittsburgh*

https://www.usenix.org/conference/soups2022/presentation/colnago

This paper is included in the Proceedings of the Eighteenth Symposium on Usable Privacy and Security (SOUPS 2022).

August 8–9, 2022 • Boston, MA, USA

978-1-939133-30-4

Open access to the Proceedings of the Eighteenth Symposium on Usable Privacy and Security is sponsored by USENIX.

Is it a concern or a preference? An investigation into the ability of privacy scales to capture and distinguish granular privacy constructs

Jessica Colnago
*Google**

Lorrie Faith Cranor, Alessandro Acquisti
Carnegie Mellon University

Kate Hazel Stanton
University of Pittsburgh

Abstract

Privacy scales are frequently used to capture survey participants' perspectives on privacy, but their utility hangs on their ability to reliably measure constructs associated with privacy. We investigate a set of common constructs (the intended objects of measurement by privacy scales) used in privacy surveys: privacy attitude, privacy preference, privacy concern, privacy expectation, privacy decision, and privacy behavior. First, we explore expert understanding of these constructs. Next, we investigate survey participants' understanding of statements used in privacy scales aimed at measuring them. We ask a balanced sample of Prolific participants in the United States to identify the extent to which different constructs describe each of a set of 30 statements drawn from scales used commonly in the privacy literature and 39 that we developed. Our analysis reveals considerable misalignment between the constructs associated with the statements and participant understanding. Many statements used in scales or that we developed with the intention to measure constructs such as privacy concern, are seen by survey participants as describing other constructs, such as privacy preferences. We also find that no statement uniquely measured any one construct, though some more reliably track their target construct than others. Our findings constitute an epistemological problem for use of scales in the existing literature (are they capturing what we think they capture?) and a practical problem for construction of new scales (how to ensure construct validity in the face of ill-defined constructs and evolving privacy landscape?). We use methods from corpus linguistics to identify characteristics of those statements most reliably associated with their target construct, and provide a set of provisional suggestions for future statement construction. Finally, we discuss the implication of our results for the privacy research community.

1 Introduction

Privacy scales are familiar instruments in privacy research [16]. These scales aim at measuring *constructs* — specific facets of participant privacy psychology, such as privacy concerns or privacy preferences — by soliciting degrees of agreement with statements believed to capture these constructs [13, 20]. A valid privacy scale can offer useful insight into public perspectives on privacy, but a scale that is not valid — that is, a scale that fails to measure its intended construct — presents a challenge for privacy research by yielding results that cannot sustain accurate generalisations and that lack predictive power [21]. Recent work has challenged the validity of existing scales [10, 18]. Here, we present evidence that problems with validity may be widespread — perhaps even intrinsic to the privacy scale as an instrument given the ill-defined and ever evolving nature of privacy — as thoroughly validated scales did not achieve conceptual clarity on the constructs they attempt to capture. We show that survey participants cannot identify *unique* constructs corresponding to statements used in scales, and that there is considerable variation in beliefs concerning which construct a statement corresponds to. There is little hope that a scale aimed at measuring, for example, *privacy concerns* can be trusted to do only that, when participants may have been understanding its constituent statements as expressing *privacy preferences*.

We investigate the following constructs, which are common in the privacy literature: *privacy attitude*, *privacy preference*, *privacy concern*, *privacy expectation*, *privacy decision*, and *privacy behavior*. Since there are no definitions of these constructs universally accepted by privacy scholars, we offered a set of definitions taken from a recent book chapter [5] to 22 privacy experts, and iteratively refined these definitions based on the experts' feedback. Next, as many privacy-related studies are performed using crowd-sourcing platforms, we

* The work was performed while Jessica Colnago was at Carnegie Mellon University.

USENIX Symposium on Usable Privacy and Security (SOUPS) 2022.
August 7–9, 2022, Boston, MA, United States.

asked a sample of Prolific participants in the United States to identify the extent to which the different constructs, presented with our refined definitions, describe each of 30 statements from scales used commonly in the privacy literature. We also asked participants to perform this task for 39 statements that reflect commonly stated privacy opinions observed in qualitative privacy studies. We leveraged Prolific's representative sample functionality to recruit a sample balanced using Census information on age, gender, and ethnicity. All studies were approved by our institution's Internal Review Board.

Our analysis shows that many statements intended to measure certain constructs that commonly appear in the privacy literature and that are systematized in Cranor and Schaub's framework [5] (for example, privacy concern) are, in fact, seen by survey participants as describing other constructs in the framework, such as privacy preferences.

We also found that no statement uniquely measured any construct. The results highlight the difficulty of using scales to measure privacy constructs uniquely and reliably. We observe that some statements were, however, more regularly matched with particular constructs. We use methods from corpus linguistics to identify features that these statements share and generalise over them to make provisional suggestions aimed at guiding future scale construction. Finally, we discuss the implication of our results for the privacy research community.

2 Background and related work

This paper builds on work in the privacy literature concerning privacy scales and privacy surveys, and on critical contributions that raise problems for those scales and surveys.

2.1 Privacy scales and privacy constructs

We focus on some of the most popular privacy scales: Westin's Privacy Segmentation Index [12], Concern for Information Privacy (CFIP) [20], Global Information Privacy Concern (GIPC) [13], and Internet Users' Information Privacy Concern (IUIPC) [13]. Some of these scales are validated—that is, carefully designed to ensure that the set of included statements consistently capture a construct. As we discuss, all of them appear to have been designed to measure *privacy concern*, as it was understood at the time of the scale's creation. We present each scale discussed in this paper and discuss how it is used in our empirical analysis. All scales are reproduced in Figure 9 in the Appendix.

Westin's Privacy Segmentation Index: Alan Westin created privacy indexes to track trends in privacy perspectives over time. Based on their answers, survey participants were classified into categories that "represent a continuum of privacy concern" [12]. To the best of our knowledge, these indexes did not form a validated scale. In particular, Westin's Privacy Segmentation Index captured participants' level of agreement on a 4-point scale to three statements. Participants who agreed with the first statement and disagreed with the second and third statements were classified as *privacy fundamentalists*. Participants who presented the opposite pattern were classified as *privacy unconcerned*. Finally, all other participants were classified as *privacy pragmatists*.

Global Information Privacy Concern: The Global Information Privacy Concern (GIPC) scale was first mentioned by Malhotra et al. in 2004 [13] and considers six statements measured on a 7-point scale. An extensive literature search for mentions of GIPC did not yield results prior to 2004. Thus, we do not have information on how these statements were selected and whether this scale has been validated. In this paper, we consider that GIPC measures concern, given the presence of this construct in the scales' name.

Concern For Information Privacy: In 1996, Smith et al. proposed the Concern For Information Privacy (CFIP) scale. This served as a first validated instrument for measuring concerns about organizational information privacy practices,but the paper does not provide a definition of concern. This scale followed a rigorous development methodology that included the generation of sample items and verification of content validity, followed by exploratory and confirmatory factor analysis, and assessments of internal validity, reliability, and generalizability. The CFIP scale includes 15 statements and four sub-scales that measure dimensions of individuals' concerns about organizational privacy practices: collection, errors, unauthorized use, and improper access. Participants report their level of agreement with each of the above statements on a 7-point scale, which are then be converted into means for the sub-scales, as well as the overall scale [20].

Internet Users' Information Privacy Concern: Malhotra et al. proposed the Internet Users' Information Privacy Concern (IUIPC) scale "[t]o reflect Internet users' concerns about information privacy" with a focus on "individuals' perceptions of fairness/justice in the context of information privacy." IUIPC was adapted from CFIP and included new items and dimensions. The authors proposed it to provide a theoretical framework on the nature of information privacy concerns for Internet users. As with CFIP, IUIPC was developed following a strict scale development methodology and results of a thorough validation process are presented in the paper. The IUIPC scale is composed of 10 statements and 3 dimensions: control, awareness, and collection (taken from CFIP). Participants report their level of agreement with each statement on a 7-point scale, and the means are calculated for each dimension [13].

2.2 Constructs and framework

We focus on a subset of constructs that have been identified to be of interest in the privacy literature: attitude, preference, con-

cern, expectation, decision, and behavior. These constructs are of long standing interest in the social sciences more broadly, where their importance and inter-relationships have been explored for decades [7]. Somewhat naturally, given such long-standing interest, we see variations in how these constructs are used across different fields, and even within the privacy literature [8]: different terms have been used to refer to the same underlying phenomenon and the same term has been used to describe slightly different phenomena over time. For example, psychologists use "worry" to refer to a "state of mental distress or agitation due to concern about an impending or anticipated event, threat, or danger" [22], while privacy scholars frequently use "concern" to refer to this state. As for "preference," the term has been used to refer to different phenomena across psychology, social sciences, and economics [11].

We leverage the conceptual framework proposed by Cranor and Schaub [5] as a seed for our construct definitions. This framework covers privacy attitude, privacy preference, privacy concern, privacy expectation, privacy decision, and privacy behavior. We used this framework due to its simplicity and coverage of central constructs used in privacy research.

2.3 Lexical issues

Constructs are specified by terms that bear rich lexical relations that complicate unique construct measurement. As noted above (see Section 2.1), and in alignment with Smith et al. [19], the privacy scales being evaluated in this paper seem to have been meant to capture privacy *concerns*. However, *concern* is a a subcategory (hyponym) of a broader class, *attitude* (hypernym) (cf. [6]). As such, any statement that falls under a subcategory (e.g. *privacy concern*) may also fall under the supercategory (*privacy attitude*), meaning that scales that claim to measure any subcategory may also be judged to measure the supercategory.

This inter-related nature of privacy constructs could explain the lack of construct validity found by previous work when investigating IUIPC [10,18]. In particular, Gross notes that the sub-scales Control and Awareness had "unsatisfactory local fit for two items ... calling the unidimensionality of these sub-scales into question" [10]. Our work builds on this past work, showing that statements used in privacy scales (as well as new statements we developed reflecting commonly stated privacy opinions) measure multiple privacy constructs, and frequently not the one originally intended.

Ambiguous or low-context statements, featured in many scales, also present problems. For example, a key difference between a *concern* and a *preference* is the affective valence of the attitude: concerns are negatively valenced whilst preferences are positively valenced. When unambiguous information about the intended affective valence is not available from the statement, this information must be supplemented by participants to determine whether the statement expresses a privacy preference or a privacy concern. For example, the statement "To me it is the most important thing to keep my privacy intact from online companies" (GIPC) may be seen as describing *privacy concern* by someone who believes corporate data collection is harmful and as describing a *privacy preference* by someone who believes corporate data collection is benign or beneficial. This supplementation may be done differently depending on individuals' priors [14].

Previous work has examined a related issue by exploring the framing of privacy-related questions [3,10]. Findings indicate that use of priming words, such as privacy or autonomy, can lead to skewed results [10]. Furthermore, it was found that surveys introduced with privacy-related warnings elicited results significantly different from those without privacy warnings [3].

A further source of complication is that statements may possess features that are connected to multiple constructs—a statement may refer both to a behavior (and so judged to measure behavior) and to negative affect (and so judged to measure privacy concern). As a result of overlapping linguistic and conceptual structures in both constructs and statements, privacy scales may be by nature unsuitable for unique construct measurement.

3 Construct definitions study

We conducted two studies to investigate the extent to which various statements regarding privacy—many of which are employed in popular privacy concern scales—are described by distinct privacy constructs: a construct definitions study with experts (discussed in this section); and a statement classification study with a balanced sample of US respondents provided by the Prolific platform (discussed in Section 4). The construct definitions study leveraged experts' opinions to define an initial set of privacy constructs and associated definitions, which we then refined through an iterative process and later provided to crowd worker participants in the statement classification study to reduce variation in interpretation of the constructs.

3.1 Methodology

In the construct definitions study, we iteratively vetted privacy constructs and definitions with privacy experts with the goal of defining a set of constructs and definitions to be used with Prolific participants in the statement classification study.

To navigate the observed variation in the literature, we first established working definitions for each construct. We started from a framework of privacy constructs and associated definitions proposed by Cranor and Schaub [5] that distinguishes privacy attitude, privacy preference, privacy concern, privacy expectation, privacy decision, and privacy behavior (Table 1). As the definitions associated with this framework had not been empirically tested, we engaged a set of privacy experts in a

Construct	Initial framework	Revised framework	Final framework
Privacy attitude	The data subject's predisposition regarding privacy, usually expressed in broad and non-actionable terms.	An individual's predisposition towards privacy (and technology) which influences their stance regarding different privacy-related situations.	An individual's predisposition towards privacy which influences their stance regarding different privacy-related situations.
Privacy preferences	What the data subject prefers to happen.	(Same as final)	An individual's preferred outcome for a specific privacy-related situation.
Privacy concern	What the data subject fears might happen.	(Same as final)	An expression of worry towards a specific privacy-related situation.
Privacy expectation	What the data subject thinks will happen.	An expression of what one views as the likely outcome of a specific privacy-related situation or behavior from the other parties involved.	An expression of what one views as the likely specific privacy-related outcome of a situation or behavior from the other parties involved.
Privacy decision	What the data subject decides or intends to do.	What an individual chooses to do in a specific situation given the resources available to support their decision making process.	What an individual chooses to do in a specific privacy-related situation among available options.
Privacy behavior	What the data subject does.	(Same as final)	What an individual actually does or has done in an attempt to achieve the level of privacy that they prefer.

Table 1: Evolution of the framework from its original format to the final version based on experts' feedback. Note that Cranor and Schaub's definition for privacy attitude was "The data subjects' general predisposition regarding privacy." We start with a modified version that the authors thought improved clarity.

process of refinement of the initial framework, so that the constructs and definitions would be generally well aligned with the experts' understanding. The refinement process took place until the feedback converged into agreement—this happened within two rounds.[1]

In the first round, we presented 22 experts (described in Section 3.2) with a survey that introduced the constructs and the initial set of associated definitions. We asked the experts whether they agreed with the definitions, and offered an open-ended response field to elaborate on points of disagreement. We also presented experts with statements from privacy scales and asked them which constructs best applied. Based on the first-round results, we generated a revised framework of constructs and associated definitions.

In the second round, we presented the revised framework to the 19 experts who had agreed to be contacted again. We received nine responses, which led to several small changes in the definitions. The initial, revised, and final iteration of the framework are shown in Table 1. In Section 3.3 we present the comments that experts provided in both the first and the second rounds of Study 1.

[1]The results of the statement classification study are robust to both providing and not providing participants with these definitions. See Section 4.

3.2 Expert selection and demographics

We selected privacy experts who worked in the areas of usable privacy, privacy law, or privacy policy; had authored at least five published papers in one of these areas in the past 10 years; and were located in the US.[2]

Two members of our research team generated an initial list of experts. We identified additional experts from the authors of papers retrieved with a search of the ACM Digital Library[3] and equivalent queries using Web of Science. After compiling a list of 68 potential experts, we verified the requirements above through online publication lists. Nine did not fit the required criteria and we could not validate nine others. Seven were not located in the US. We contacted the remaining 43 experts via email. We obtained complete responses from 22 experts in round 1 and 9 experts in round 2.

In the first round, half of the experts self-identified as male and half as female. On average, the experts had 16 years of experience with privacy research (sd: 5.9 years). In the second round, three self-identified as male, and six as female. On average, the experts had 16 years of experience with privacy research (sd: 8.9 years). In the first round, 11 experts

[2]This was a requirement of our Internal Review Board due to concerns about the General Data Protection Regulation that they had not resolved at the time of our study.

[3]Search Queries: [All: "privacy policy"] OR [All: "privacy law"] OR [All: "usable privacy"] AND [Publication Date: (01/01/2010 TO 01/31/2020)]; and analogous searches with only one research area at a time

described their background as "Social Sciences," nine "Computer Science," five "Law," and three "Other." The majority of the experts reported working in academia, with two citing industry experience, and one mentioning policy and government. Only one expert listed industry and only one expert listed policy as their main area of focus. The second round had a mix of law, computer science, and social science in a similar proportion as the first round.

3.3 Expert feedback on definitions

The first round of feedback highlighted experts' concerns over lack of clarity of some definitions. Some comments were targeted at the vagueness of the initial set of definitions: "The description lacks an indication of what the preference is about." Others addressed specific word choices: "I am not sure that concern = fear. One can have legitimate concerns without being fearful." Some experts suggested that we better tie the definitions to privacy: "The definition would need to be completed by indicating 'what the data subject does with respect to privacy'." This initial round of feedback led to significant changes to the initial set of definitions, as can be seen in Table 1. The revised set was presented again to experts in the second round of the study.

The second round of feedback was narrower and pointed, leading to the final framework presented in Table 1. Below, we summarize the feedback we received in the second round.

Privacy attitude: One expert pointed out that a parenthetical in "predisposition towards privacy (and technology) ..." could be confusing. We agreed and removed the parenthetical. Another expert asked whether the definition only applied to attitudes about one's self, or if it also applies for attitudes towards others (for example, "I think my kids should be more careful sharing information on Facebook"). We decided that the existing definition appropriately included both and did not revise further.

Privacy preference: In the second round we did not receive any feedback for this construct.

Privacy concern: One expert highlighted that there may be a fundamental difficulty with measuring concern, as concern is a combination of expectation and trust. One may not be concerned about an otherwise concerning issue because they trust the parties involved. While we agree, as our focus is not on sources of concern, we did not revise the definition.

Privacy expectation: One expert noted that the phrasing of the definition suggests that all outcomes of a privacy-related situation are privacy expectations, even if some are not related to privacy. We reworded so "privacy-related" modifies "outcome" rather than "situation."

Privacy decision: An expert pointed out that our definition did not mention privacy. We revised our definition to refer to decisions in "privacy-related" situations and added that a decision can only be made from a set of available options.

Privacy behavior: This construct received the strongest negative review, with one expert stating:

> This definition I disagree most with – I think privacy behaviors are often inconsistent with what people would prefer and many behaviors are in conflict with the level of privacy that people prefer. I think privacy behavior is what an individual does that has an impact on their privacy, regardless of whether it's positive or negative or consistent with their attitudes, preferences, or concerns.

While we agree that privacy behaviors may not always achieve a person's desired outcome and may even be counter productive, we think it is important to limit this definition to behaviors that were intended to achieve a privacy-related outcome. For example, while closing curtains is a behavior that can increase privacy, people also close curtains for other reasons, such as reducing screen glare or darkening a room. For this reason, it is important that behavior-related statements specify the goal of said behavior.

4 Statement classification study

The statement classification study used data from online crowd worker participants—a typical population of focus for measuring privacy perspectives—to assess which constructs and definitions defined in the construct definitions study described a set of 69 privacy statements. We took 30 statements from existing privacy scales and developed 39 additional statements. For each of the new statements we developed, we classified it according to the authors' expectations as to the construct with which it would best align.

We presented participants with the following prompt: "Imagine that you are talking to a friend, and your friend says the following sentence." This was followed by a randomly selected statement. We asked participants to rate how well each of the constructs described what their friend was saying in that sentence. Participants rated each construct on a 5-point scale, from "Does not describe at all" (1) to "Describes very well" (5). Each participant was presented with a random selection of seven statements out of the 69 available. Each statement was rated by approximately 40 participants.

Since, in pilot studies, we did not identify differences in how participants classified statements between the group that was shown the constructs with the definitions and the one that only saw the constructs, and given our desire to normalize participants' interpretations of the constructs to the maximum possible extent, we showed all participants the constructs and associated definition for each classification task.

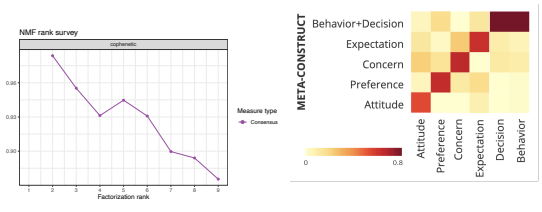

Figure 1: Left: Cophenetic correlation coefficient graph (ranks 2–10) showing a continuous drop for ranks >5. Right: NMF basis results showing the composition of the meta-constructs. Values were normalized to range from 0 to 1.

4.1 Participant Demographics

We recruited 400 participants from the Prolific platform. Prolific's representative US sample provides a balanced sample in terms of gender, age, and ethnicity based on US Census data. Fifty percent in our sample self-identified as female, with one participant choosing non-binary. The mean age was 46.4 years, with a standard deviation of 16.3 years. When asked about their ethnicity, 71% of our participants self-identified as white, 14% as Black or African American, 8% as Asian, 6% as Other (which could encompass mixed race), and one participant self-identified as American Indian or Alaska Native. Furthermore, 7% of our participants self-identified as of Hispanic or Latinx. Lastly, 16.5% of our participants reported working in or studying a technology related area.

4.2 Analyses Approach

We first binned participant responses for every statement into "high" (4 or 5) and "low" (1, 2 or 3) scores. To check the robustness of this approach, we compared results when binning the neutral option (3) with both the high and low categories. The differences observed did not impact the findings we present.

For each statement we determined whether there was a "primary construct" as follows. We identified the two constructs with the highest count of high scores (from approximately 40 responses) and compared their counts of high and low scores. We used Chi-square tests and Cramer's V to determine whether the top construct was statistically different from the second highest one. The distributions were considered distinct if the p-value from the Chi-square was smaller than 0.05; otherwise, they were considered similar. For distinct distributions we report the effect size using Cramer's V. The results are presented in Section 4.3.1.

The results of this analysis indicated that the majority of statements were not described by a single primary construct, and that those that were often had small effect sizes. Therefore, we turned next to an analysis approach that did not rely on distinct constructs and could provide insights into how the constructs related to one another. We used Non-negative Matrix Factorization (NMF) which automatically "extract[s] sparse and easily interpretable factors" [9]. This method provides a better understanding on how the constructs relate to one another and how they relate to the statements. We ran the algorithm on a matrix composed of the six constructs and 69 statements. Each cell corresponded to the count of participants who selected a "high" level of agreement (Strongly agree (5) or agree (4)) for each construct statement pair.

Similar to cluster analysis, the first step in NMF is to identify how many ranks, similar to groups and clusters, will lead to stable and descriptive results. While there are many ways of selecting the rank [9], in this work we do so by examining the cophenetic correlation coefficient graph (Figure 1, left) obtained from the consensus matrix—the average connectivity matrix over many clustering runs [4].[4]

Following the rule of "select[ing] values of k where the magnitude of the cophenetic correlation coefficient begins to fall" [4], we selected five ranks, for which the algorithm outputs five basis components—we refer to these components as "meta-constructs." These meta-constructs are a composition of the initial constructs and, as we can see in Figure 1 (right), they roughly break along the lines of the constructs, with privacy behavior and privacy decision being grouped in a single meta-construct.

By using the consensus output obtained from running the algorithm 100 times, the NMF algorithm associates each statement with a meta-construct. Thus we produced five groups of statements corresponding to our meta-constructs. We present our results in Section 4.3.2.

4.3 Statement classification results

We present our classification results based on primary constructs and meta-constructs, as well as broken down by scale.

4.3.1 Primary constructs

We see that only 33 of the statements (48%) had the top construct statistically different from the second highest one. This means that there was a primary construct that survey participants perceived as describing individual statements for roughly half of the statements. Even among those, none had a large effect size: 23 had a low effect size ([0.1, 0.3]) and ten had medium effect sizes ([0.3, 0.5]). For the rest, no primary construct was identified. The right side of Figures 2 through 6 show the percentage of high selections in green, highlighting those that had a primary construct with a dotted box.

4.3.2 Meta-constructs

Our findings for primary constructs seem to indicate a lack of independence between the constructs and definitions that we used. Therefore, we used NMF to identify composite

[4]The consensus matrix was obtained through 100 iterations of the algorithm.

Construct	CFIP	GIPC	IUIPC	Westin	New
Attitude	0	2	2	1	10
Preference	0	1	1	0	4
Concern	2	1	0	0	5
Expectation	0	0	0	0	1
Decision	0	0	0	0	3
Behavior	0	0	0	0	0

Table 2: Breakdown of the number of statements with primary constructs per source. For IUIPC we only consider the six statements unique to IUIPC, those related to Control and Awareness.

Construct	CFIP	GIPC	IUIPC	Westin	Self-gen
Attitude	0	2	1	0	8
Preference	6	1	3	0	6
Concern	6	2	1	1	6
Expectation	2	0	1	2	6
D & B	1	1	0	0	13

Table 3: Breakdown of the number of statements each meta-constructs per source. For IUIPC we only consider the six statements unique to IUIPC.

constructs. The NMF results show the weighted function of the identified meta-constructs that describes each statement (see left heatmap on Figures 2 through 6).

4.3.3 Results by scale

We present our results with statements grouped according to the scale in which they are used. Tables 2 and 3 summarizes the breakdown of primary constructs and meta-constructs by source. Figures 2 through 6 also include the scale for each statement and the construct to which the scale authors expected or intended it to align.

CFIP: This scale was intended to capture the construct *privacy concern*. Out of the 15 statements that compose CFIP, we found that only six had *privacy concern* as their meta-construct (Figure 4), while six others had *privacy preference* as their meta-construct. Of note, "Companies should have better procedures to correct errors in personal information" and "Companies should take more steps to make sure that the personal information in their files is accurate" were associated with the meta-construct *privacy expectation*, though Figure 5 shows that none of the meta-constructs seem to be dominant.

GIPC: While we could not establish it with certainty, we consider that the underlying construct intended to be measured by GIPC's statements is *privacy concern*. We see a similar pattern to CFIP, where GIPC's statements were infrequently associated with *privacy concern* as their meta-construct. Two out of the six GIPC statements had *privacy concern* as their meta-construct. Interestingly, the statements "I believe other people are too much concerned with online privacy issues" and "Compared with other subjects on my mind, personal privacy is very important" had *privacy attitude* as their meta-construct.

IUIPC: We consider that IUIPC had the intention to capture the construct *privacy concern*. For the six statements related to awareness and control, which were created for IUIPC, we see that *privacy concern* was the meta-construct for only one statement: "I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction." Instead, three statements had *privacy preference* as their meta-construct. "Consumer control of personal information lies at the heart of consumer privacy" had *privacy attitude* as its meta-construct while "It is very important to me that I am aware and knowledgeable about how my personal information will be used" had *privacy expectation*.

Westin: We consider that Westin's Privacy Segmentation Index statements were created with the intent to measure *privacy concern*. However, what we found is a combination of concern and expectation. The statement "Consumers have lost all control over how personal information is collected and used by companies" had *privacy concern* as its meta-construct, though attitude was more frequently selected. The statements "Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today" and "Most businesses handle the personal information they collect about consumers in a proper and confidential way" had *privacy expectation* as their meta-construct.

Generated statements: We also examined the statements that we generated for the study, considering our specific constructs and definitions. Our expected construct matched the meta-construct predominantly selected as describing the statement for about 85% of the statements. As we can see in the heatmap figures, the statements that did not match were:

- I am not satisfied with my current level of privacy (Expected: attitude; classification: concern)
- I don't care about privacy as long as I can use the service (Expected: preference; classification: behavior and decision)
- I don't think there's anything to worry about privacy (Expected: concern; classification: attitude)
- I will be able to achieve the level of privacy that I want to have (Expected: expectation; classification: preference)

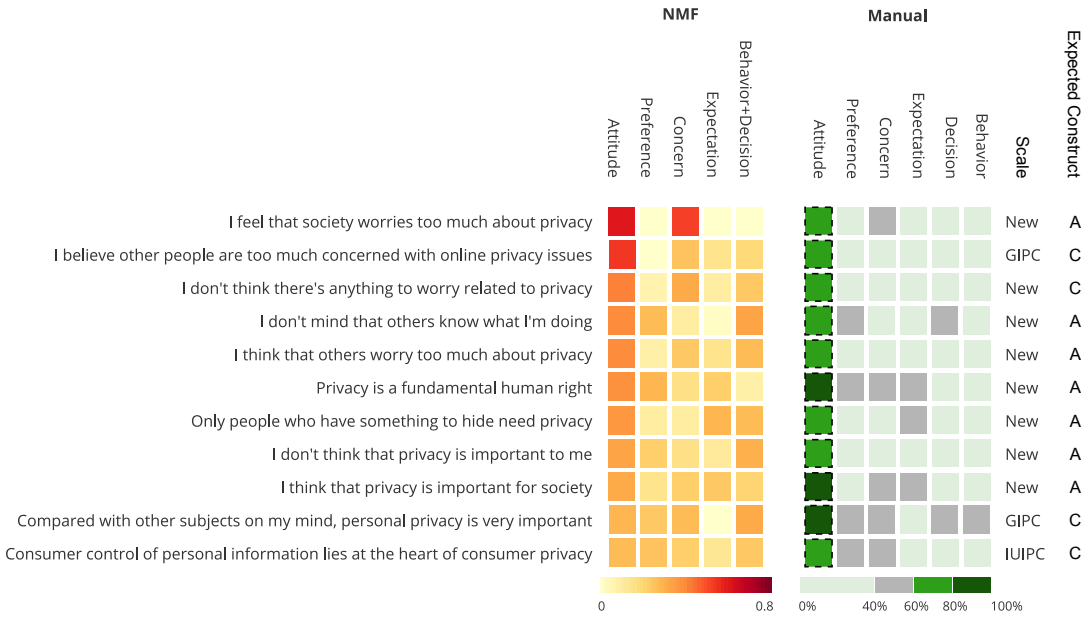

Figure 2: Heatmap displaying the NMF coefficient results showing the composition of each statement based on the meta-constructs (left) and the percentages of high scores for each construct/statement pair (right) for statements under the "attitude" meta-construct. The primary construct identified is highlighted by a dotted box.

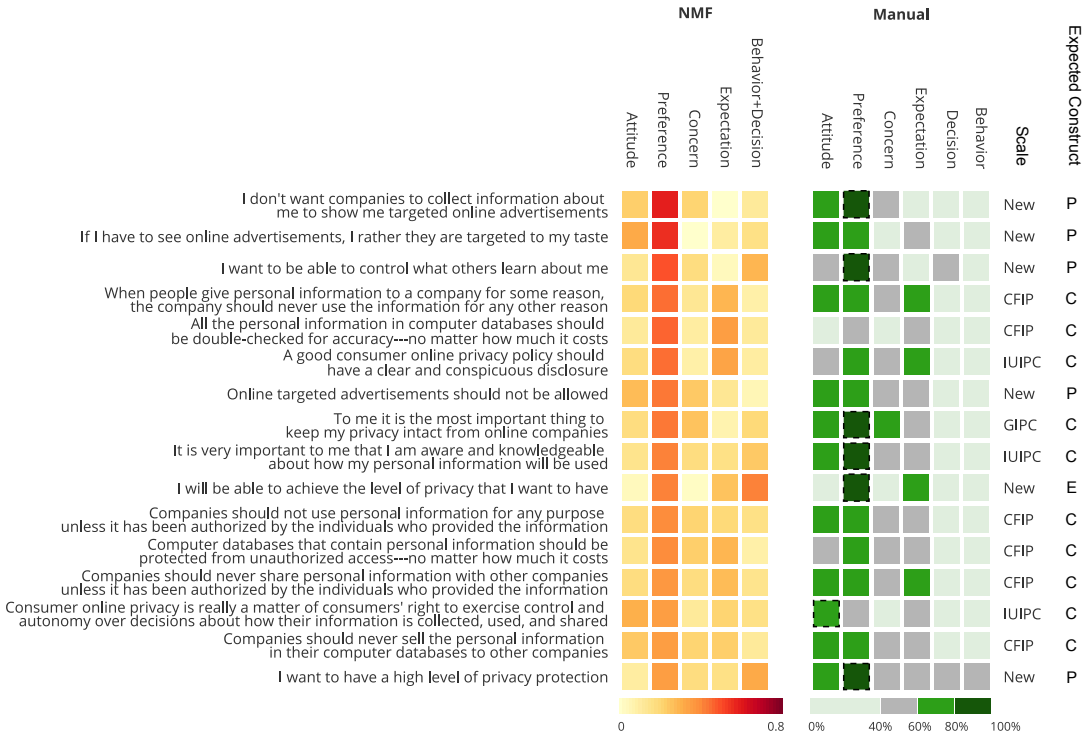

Figure 3: Heatmap displaying the NMF coefficient results showing the composition of each statement based on the meta-constructs (left) and the percentages of high scores for each construct/statement pair (right) for statements under the "preference" meta-construct. The primary construct identified is highlighted by a dotted box.

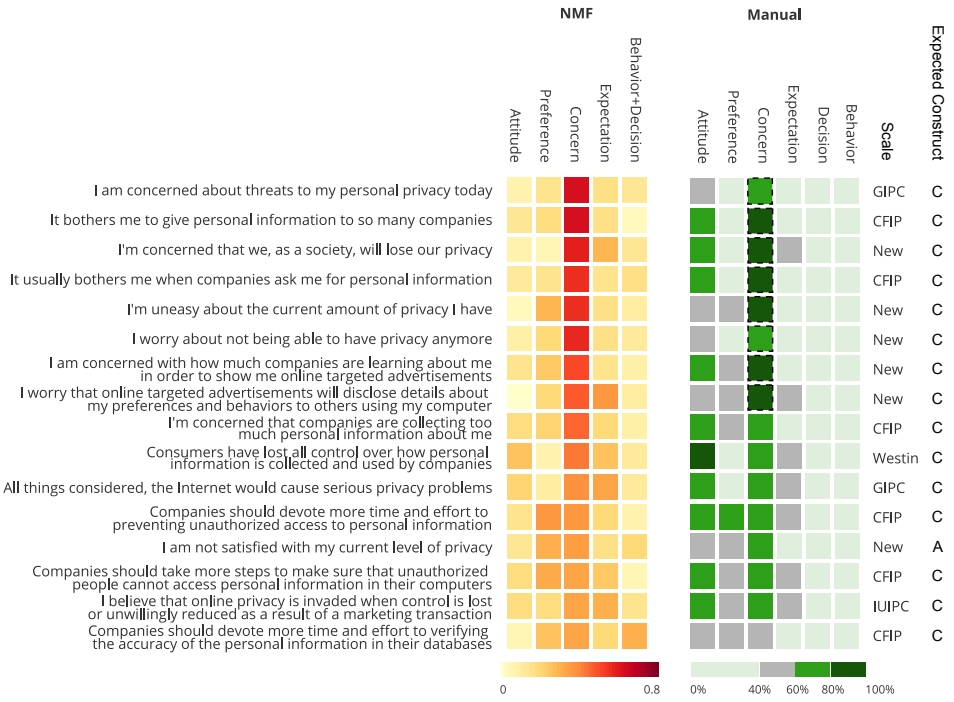

Figure 4: Heatmap displaying the NMF coefficient results showing the composition of each statement based on the meta-constructs (left) and the percentages of high scores for each construct/statement pair (right) for statements under the "concern" meta-construct. The primary construct identified is highlighted by a dotted box.

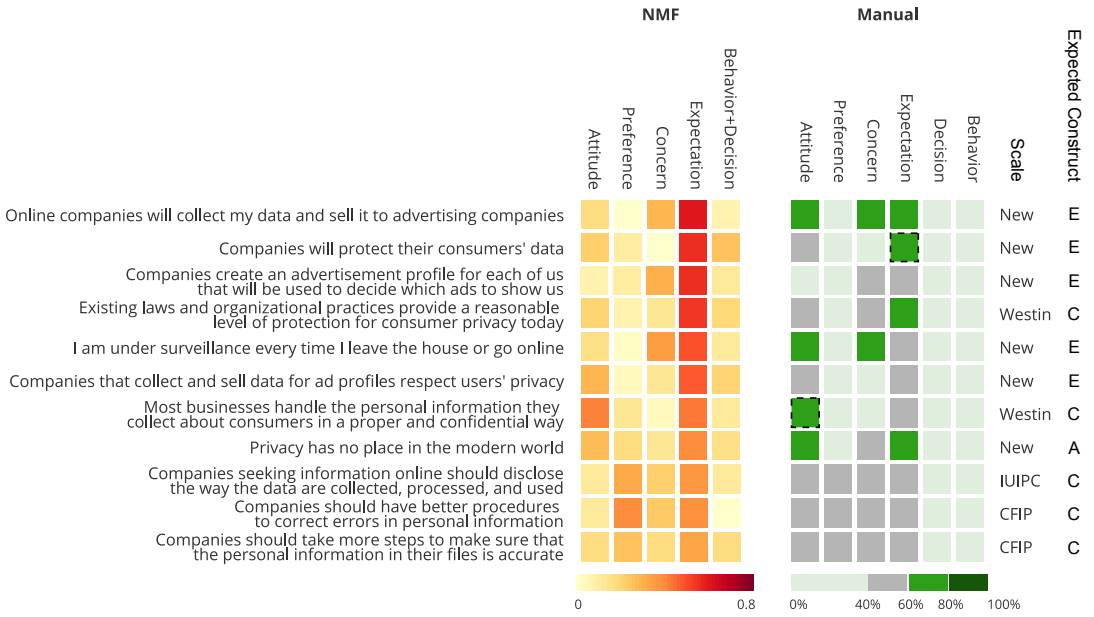

Figure 5: Heatmap displaying the NMF coefficient results showing the composition of each statement based on the meta-constructs (left) and the percentages of high scores for each construct/statement pair (right) for statements under the "expectation" meta-construct. The primary construct identified is highlighted by a dotted box.

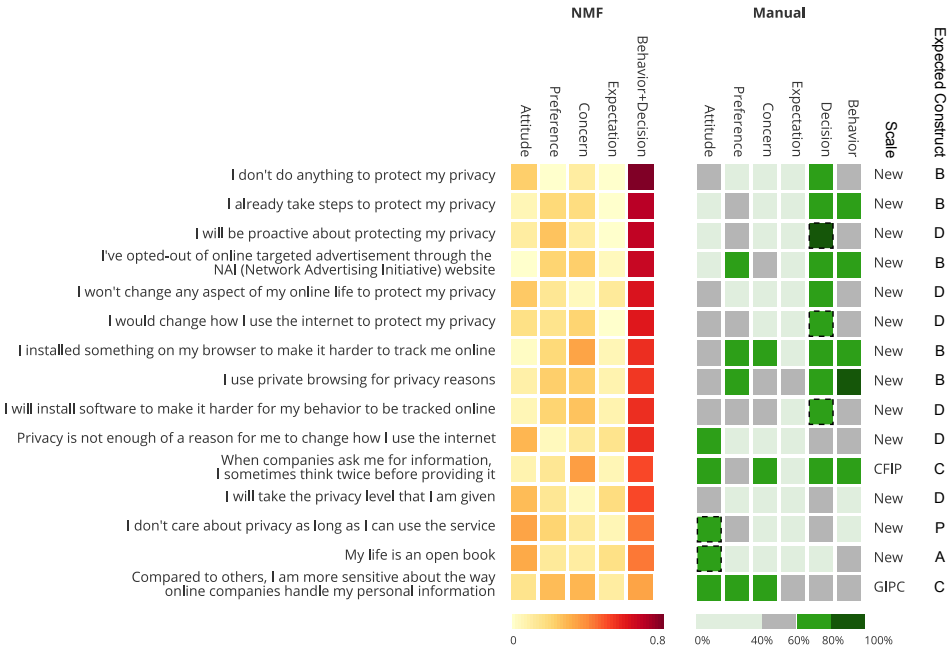

Figure 6: Heatmap displaying the NMF coefficient results showing the composition of each statement based on the meta-constructs (left) and the percentages of high scores for each construct/statement pair (right) for statements under the "decision-behavior" meta-construct. The primary construct identified is highlighted by a dotted box.

- My life is an open book (Expected: attitude; classification: behavior and decision)
- Privacy has no place in the modern world (Expected: attitude; classification: expectation)

This suggests that even when building statements with specific constructs in mind, misalignment occurs between researchers' goals and survey participants' interpretations. In the next section we examine some of the linguistic patterns used in these statements that tend to be problematic or that tend to be associated with particular constructs. An awareness of these patterns may help researchers write statements that will be more likely to be interpreted as intended.

4.4 Corpus analyses on NMF groups

We conducted a corpus analysis to investigate whether linguistic patterns could be found that might help minimize problematic conceptual and lexical overlaps. The findings presented in Section 4 showed that some statements may be more strongly correlated with particular constructs; any regularities in the kinds of expression that occur in those cases could potentially be exploited in scale construction to improve researcher control over which constructs are being measured.

We constructed corpora (sets of statements) from the groups derived from NMF analysis. These were then analysed using Wmatrix [17]. WMatrix assigns broad semantic field categories and calculates overuse and under use of semantic field categories between corpora. The software compares relative frequencies within the data and calculates log-likelihood and log ratio. We compared between construct corpora and the AMe06 corpus of written, published, American usage [15]. We discuss selected results of log likelihood analysis.[5] High log likelihood ($p < 0.001 - p < 0.05$) represents statistically significant overuse of a semantic field in NMF corpus relative to AME06.[6] Table 5 in the Appendix displays the binary log of the ratio of relative frequencies (log ratio) across statistically significant categories.

The following general patterns provide an instructive start. The *privacy attitude* corpus significantly overrepresented a range of semantic categories that unambiguously signal that the speaker is expressing an attitude or making an evaluation. Attitude verbs, nouns relating mental or conceptual objects, such as **thought**; **comparative judgements** and **judgments of importance** were prevalent in statements strongly correlated with *privacy attitude*. As noted above, 'concern', and 'preference' are sometimes considered subcategories of 'attitude' and so overlaps in overrepresentation were to be anticipated and were found; expressions signalling worry were overrepresented in both the *privacy attitude* corpus, and the

[5] See Appendix for full table of log ratio analyses. Log ratio is a metric of effect size, each point reflecting a doubling of the rate of occurrence in the NMF corpora relative to the AME06

[6] Unsurprisingly, given the context, certain categories (e.g. *Information technology and computing*; *business: generally*; *business: selling*) are overrepresented across the corpora. These categories are common thematic topics across corpora.

privacy concern corpus, and value judgment categories occurred in both *privacy attitude* and *privacy preference*. However, the *privacy preference* corpus distinguished itself by overrepresentation of **verbs signalling desire and modals signalling desired outcomes**, including 'want' (under *Wanted*), 'should' (under *Strong obligation or necessity*) and 'never' (under *Time*). *Privacy concern* corpus distinguished itself with over-representation of a range of expressions signalling **negative attitudinal valence**, including attitude verbs and deverbal expressions, as seen under the categories *Worry* and *Failure*, along with negative morphemes (e.g. 'un' in 'unauthorized').

The *privacy expectation* corpus over-represented **future auxiliaries**, for example, 'will' under *Time: future* — a category also overrepresented in *privacy decision and behavior*. It was distinguished from the latter, however, by overrepresentation of **value judgments**. The decision and behavior corpus distinguished itself in over-representation of a range of **privacy-behaviour related verbs** (in categories: *Helping* (mainly populated by 'protect') and *Investigate, examine, test, search*) and verbs with privacy-related direct objects.

Perhaps the primary lesson to be extracted from this analysis is that statement interpretation is considerably more open-ended than has been previously accounted for. This open-endedness may be to some extent ineliminable due to close relations between the constructs.

Statements that saw least convergence between participants were long or syntactically complex — both factors increase the potential for participants to draw on distinct information sources leading to diverging interpretations. Shorter affectively ambiguous declaratives (i.e. declaratives with no clear indication of whether the content is intended to describe a positive or negative state of affairs) also led to high variation by participants, since lack of information leads to speakers supplementing background beliefs to extract an interpretation.

Those statements that saw greatest convergence between participants on a particular construct, suffered neither from excess length or brevity and bore features that encouraged participants to navigate the possibilities in similar ways. Statements aimed at measuring constructs signalling attitude types, for example, can be improved by including attitude verbs that clearly signal those types (for concern, 'I worry/fear/am concerned that' for preference 'I like/prefer that/am comfortable with'). These provisional suggestions are not, however, programmatic, and should rather highlight work to be done in isolating linguistic factors that could help constrain participant interpretation.

5 Limitations

Our results are limited by a number of factors.

Sample: While we attempted to produce results that could be generalizable to the sample populations typically used in privacy studies by leveraging Prolific's representative sample, our results may still not generalize beyond that sample.

Analysis approach: While NMF is, to the best of our knowledge, the most well-suited method for the problem at hand, the algorithm may yield slightly different results in different executions. We minimized this by leveraging best practices, such as performing multiple executions and utilizing the consensus results. In our executions of the algorithm, these variations did not impact the findings presented here. Furthermore, our results are limited by the threshold selected for our analyses. We minimized potential issues with threshold selection by performing robustness checks, finding no significant impact to the findings.

Definitions: The definitions we proposed are a best-effort at an initial set to be used by the privacy community. However, they still need to be improved and more broadly vetted. Furthermore, while we tried to reduce the variation in interpretation of the constructs by providing participants with the associated definitions, there are no guarantees that the definitions were interpreted in the same manner by all participants.

6 Discussion

We presented the results from an investigation of constructs captured in privacy scales. First, we refined a set of definitions for commonly used privacy constructs with the aid of privacy experts. Next, we used these definitions to collect participants' views on which constructs describe each of 69 statements. Those statements represent a collection of both newly generated statements and statements from privacy scales.

Our results suggest that statements from existing privacy scales measure multiple constructs simultaneously, and often represent constructs other than concern, which appears to be the intended construct. To a lesser degree, a similar phenomenon happens with statements that were designed with the constructs in mind. The observed lack of a one-to-one match between statement and construct is, arguably, a result of two separate factors: the inherent ambiguity of natural language and the overlap between privacy constructs. The observed mismatch between statements and constructs may be due in part to a lack of agreed upon definitions for different privacy constructs, and on the evolving understanding [1] and use of these terms since the scales' creation.

We show that is is possible to leverage aspects of semantics and sentence structure to help participants identify a target construct. In general, simpler sentences that provide sufficient information to the reader, so that their range of interpretation is reduced, seem to be more successful at reducing variation in interpretation. Nevertheless, we must be mindful of how this information is framed to avoid eliciting an exaggerated emotional response [3, 10].

Nevertheless, it may be ultimately unlikely that we can create *statements* that *only* measure a specific construct. In this

paper, we show that the constructs considered in the privacy community are not perceived as fully independent—attitude, preference, concern, and expectation were frequently simultaneously selected, and behavior and decision were always simultaneously selected. This overlap between constructs likely explains why we, and previous work [10], observed how validated scales such as CFIP and IUIPC, which have shown high internal validity, contain statements that were described by multiple constructs: existing scales seem to be measuring a higher level construct, such as privacy *perspective*. Given that existing scales do not seem to uniquely measure the finer grained constructs the community commonly uses, as they are currently understood, moving forward we should acknowledge this issue and consider its impact on results.

Narrow interpretations based on the outputs of such scales and related statements have led to inconsistent findings such as the privacy paradox [1, 2, 8]. In addition to the many explanations already found for the paradox, fundamental issues may exist with the construct validity of our measuring tools.

7 Future work

There are different approaches that the privacy community can take in face of these results. Here we list a few possibilities, but they are not meant to be prescriptive or comprehensive.

Shared definitions: In this paper we present a set of definitions constructed with the aid of a diverse sample of privacy experts in the field. However, we acknowledge that this set does not necessarily have to be the one we agree to use as a community. Going forward *we need to discuss what these, and potentially other, constructs mean and develop a shared and consistent vocabulary.*

Scale development: The results presented under Section 4.4 could help in the creation of scale statements. Nevertheless, future efforts in developing scales *should take care in acknowledging the inherent and possibly systemic limitations of such tools within the privacy context.* In particular, these efforts should validate that the developed scale actually measures the construct it claims to measure and that, in all likelihood, the scale will measure a combination of related constructs. Furthermore, we *should conduct periodic assessments to ensure that scales are still in alignment with the contemporaneous understanding of these constructs.*

Measuring granular constructs: Given the overlap between more granular privacy-related constructs and the contextual nature of privacy, it is worth considering alternate methods of capturing these constructs beyond static, validated scales. *If a distinction between constructs is important to the research question at hand, using methods that allow researchers to follow up and tease apart the differences between constructs might be necessary.* For example, to distinguish preferences, concerns, and expectations, participants might be be given a description of a type of data collection and asked whether they would prefer to allow or restrict it from happening with their data (preference), whether they are worried about it happening (concern), and whether they believe it is happening (expectation).

8 Conclusion

We presented research meant to investigate our ability to uniquely and reliably capture people's granular privacy perspectives. In particular, we focus on privacy attitude, preference, concern, expectation, decision, and behavior.

We found that existing, and newly developed, statements meant to capture specific privacy constructs frequently capture multiple constructs at once. This enmeshed nature of the explored privacy constructs could help explain why existing scales, while thoroughly validated when proposed, do not always succeed at providing predictive insights, for example, as to people's engagement with privacy behaviors based on their privacy concerns. As an aid to future work developing privacy scales, we present key linguistic characteristics that could help in the creation of statements that more uniquely discern between constructs.

We further propose that future work create a well-accepted set of definitions for privacy constructs; take into account the limitations of existing privacy scales when leveraging them; periodically verify the alignment between scales and the contemporaneous understanding of what they are meant to capture; and, be mindful of the enmeshed nature of these privacy constructs, using appropriate research methods to tease them apart, when needed.

Acknowledgements

This work was supported in part by gifts from NortonLifeLock, Google, Innovators Network Foundation, and the Carnegie Corporation of New York.

References

[1] Alessandro Acquisti, Laura Brandimarte, and Jeff Hancock. How privacy's past may shape its future. *Science*, 375(6578):270–272, 2022.

[2] Alessandro Acquisti, Laura Brandimarte, and George Loewenstein. Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology*, 30(4):736–758, 2020.

[3] Alex Braunstein, Laura Granka, and Jessica Staddon. Indirect content privacy surveys: Measuring privacy with-

out asking about it. *SOUPS 2011 - Proceedings of the 7th Symposium on Usable Privacy and Security*, 2011.

[4] Jean-Philippe Brunet, Pablo Tamayo, Todd R. Golub, and Jill P. Mesirov. Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences*, 101(12):4164–4169, 2004.

[5] Lorrie Faith Cranor and Florian Schaub. Usable and Useful Privacy Interfaces. In *An Introduction to Privacy for Technology Professionals, Second Edition*, chapter Chapter 5. IAPP, 2020.

[6] D. Alan Cruse. *Hyponymy and Its Varieties*, pages 3–21. Springer Netherlands, Dordrecht, 2002.

[7] Martin Fishbein and Icek Ajzen. Introduction. In *Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research*, chapter Chapter 1. Addison-Wesley, 1975.

[8] Nina Gerber, Paul Gerber, and Melanie Volkamer. Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers and Security*, 77:226–261, aug 2018.

[9] Nicolas Gillis. The why and how of nonnegative matrix factorization, 2014.

[10] Thomas Groß. Validity and reliability of the scale internet users' information privacy concerns (iuipc). *Proceedings on Privacy Enhancing Technologies*, 2021(2):235–258, 2021.

[11] Sven Ove Hansson and Till Grüne-Yanoff. Preferences. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, summer 2018 edition, 2018.

[12] Harris Interactive. Privacy on and off the internet: What consumers want. Technical report, Harris Interactive Inc, 2002.

[13] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research*, 15(4):336–355, dec 2004.

[14] Eric McCready. Emotive equilibria. *Linguistics and Philosophy*, 35, 05 2012.

[15] Amanda Potts and Paul Baker. Does semantic tagging identify cultural change in british and american english? *International Journal of Corpus Linguistics*, 17, 12 2012.

[16] Sören Preibusch. Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human Computer Studies*, 71(12):1133–1143, 2013.

[17] Paul Rayson. From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4):519–549, 2008.

[18] Janice C Sipior, Burke T Ward, and Regina Connolly. Empirically assessing the continued applicability of the iuipc construct. *Journal of Enterprise Information Management*, 2013.

[19] H. Jeff Smith, Tamara Dinev, and Heng Xu. Information Privacy Research: An Interdisciplinary review. *MIS Quaterly*, 35(4):1689–989—-1015, 2011.

[20] H. Jeff Smith, Sandra J. Milberg, and Sandra J. Burke. Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, 20(2):167, jun 1996.

[21] Allison Woodruff, Vasyl Pihur, Sunny Consolvo, Laura Brandimarte, and Alessandro Acquisti. Would a privacy fundamentalist sell their DNA for $1000... if nothing bad happened as a result? The Westin categories, behavioral intentions, and consequences. *SOUPS '14: Proceedings of the Tenth Symposium On Usable Privacy and Security*, pages 1–18, 2014.

[22] Worry. *APA Dictionary of Psychology*. American Psychological Association.

9 Appendix

Westin's Privacy Segmentation Index

- Consumers have lost all control over how personal information is collected and used by companies.
- Most businesses handle the personal information they collect about consumers in a proper and confidential way.
- Existing laws and organizational practices provide a reasonable level of protection for consumer privacy today

GIPC

- To me it is the most important thing to keep my privacy intact from online companies.
- Compared with other subjects on my mind, personal privacy is very important
- Compared to others, I am more sensitive about the way online companies handle my personal information
- I believe other people are too much concerned with online privacy issues.
- I am concerned about threats to my personal privacy today.
- All things considered, the Internet would cause serious privacy problems

CFIP

Errors

- All the personal information in computer databases should be double-checked for accuracy---no matter how much it costs.
- Companies should have better procedures to correct errors in personal information.
- Companies should devote more time and effort to verifying the accuracy of the personal information in their databases.
- Companies should take more steps to make sure that the personal information in their files is accurate.

Unauthorized use

- Companies should not use personal information for any purpose unless it has been authorized by the individuals who provided the information.
- When people give personal information to a company for some reason, the company should never use the information for any other reason.
- Companies should never share personal information with other companies unless it has been authorized by the individuals who provided the information.
- Consumer online privacy is really a matter of consumers' right to exercise control and autonomy over decisions about how their information is collected, used, and shared.

Improper access

- Companies should devote more time and effort to preventing unauthorized access to personal information.
- Computer databases that contain personal information should be protected from unauthorized access---no matter how much it costs.
- Companies should take more steps to make sure that unauthorized people cannot access personal information in their computers.

IUIPC

Awareness

- Companies seeking information online should disclose the way the data are collected, processed, and used.
- A good consumer online privacy policy should have a clear and conspicuous disclosure.
- Companies should never sell the personal information in their computer databases to other companies.
- It is very important to me that I am aware and knowledgeable about how my personal information will be used.

Control

- I believe that online privacy is invaded when control is lost or unwillingly reduced as a result of a marketing transaction.
- Consumer control of personal information lies at the heart of consumer privacy.
- Consumer online privacy is really a matter of consumers' right to exercise control and autonomy over decisions about how their information is collected, used, and shared.

Collection

(Used in both IUIPC and CFIP)

- It usually bothers me when (online) companies ask me for personal information.
- When (online) companies ask me for information, I sometimes think twice before providing it.
- It bothers me to give personal information to so many (online) companies.
- I'm concerned that (online) companies are collecting too much personal information about me.

Figure 7: Statements for each of the scales evaluated in this paper.

#	Statement
1	Companies create an advertisement profile for each of us that will be used to decide which ads to show us.
2	Companies that collect and sell data for ad profiles respect users' privacy.
3	Companies will protect their consumers' data
4	I already take steps to protect my privacy
5	I am concerned with how much companies are learning about me in order to show me online targeted advertisements.
6	I am not satisfied with my current level of privacy
7	I am under surveillance every time I leave the house or go online.
8	I don't care about privacy as long as I can use the service
9	I don't do anything to protect my privacy.
10	I don't mind that others know what I'm doing
11	I don't think that privacy is important to me
12	I don't think there's anything to worry related to privacy.
13	I don't want companies to collect information about me to show me targeted online advertisements.
14	I feel that society worries too much about privacy
15	I installed something on my browser to make it harder to track me online
16	I think that others worry too much about privacy
17	I think that privacy is important for society
18	I use private browsing for privacy reasons
19	I want to be able to control what others learn about me
20	I want to have a high level of privacy protection.
21	I will be able to achieve the level of privacy that I want to have.
22	I will be proactive about protecting my privacy.
23	I will install software to make it harder for my behavior to be tracked online.
24	I will take the privacy level that I am given.
25	I won't change any aspect of my online life to protect my privacy.
26	I worry about not being able to have privacy anymore.
27	I worry that online targeted advertisements will disclose details about my preferences and behaviors to others using my computer.
28	I would change how I use the internet to protect my privacy.
29	I'm concerned that we, as a society, will lose our privacy.
30	I'm uneasy about the current amount of privacy I have.
31	I've opted-out of online targeted advertisement through the NAI (Network Advertising Initiative) website.
32	If I have to see online advertisements, I rather they are targeted to my taste.
33	My life is an open book.
34	Online companies will collect my data and sell it to advertising companies.
35	Online targeted advertisements should not be allowed.
36	Only people who have something to hide need privacy.
37	Privacy has no place in the modern world.
38	Privacy is a fundamental human right.
39	Privacy is not enough of a reason for me to change how I use the Internet.

Table 4: List of candidate statements created for the purpose of this study.

Att	Pref	Conc	Exp	Beh+Dec	Category
–	2.81	–	–	–	Able/intelligent
–	–	–	–	3.53	Alive
3.38	3.42	–	–	–	Allowed
3.23	4.03	4.18	4.54	3.01	Business: generally
3.73	3.35	3.00	4.55	–	Business: selling
4.97	3.46	4.27	3.60	5.02	Closed; hiding/hidden
1.85	–	–	–	–	Comparing: different
4.62	–	–	–	4.81	Comparing: similar
–	10.21	–	–	–	Double-check
4.43	–	–	–	–	Exceed; waste
3.78	–	4.32	–	–	Failure
–	–	–	1.49	–	General actions / making
2.08	1.51	–	2.59	–	Getting and possession
–	–	–	–	3.01	Helping
2.75	–	–	–	–	Important
4.10	4.46	4.05	4.02	5.03	Information technology and computing
–	–	–	–	2.92	Investigate, examine, test, search
–	5.06	–	6.35	–	Knowledge
2.45	3.56	2.89	2.22	–	Knowledgeable
–	3.72	–	–	–	Learning
–	–	4.87	–	–	Like
2.26	–	–	–	–	Mental object: conceptual object
–	2.77	–	–	–	Money: cost and price
–	–	5.33	–	–	Not allowed
4.60	5.04	5.56	4.11	4.38	Not part of a group
–	–	–	–	1.08	Pronouns
–	3.32	–	–	–	Reciprocal
–	–	–	5.17	–	Sensible
–	3.00	–	–	–	Strong obligation or necessity
–	–	–	–	4.06	Texture
–	2.64	–	–	–	Time
–	–	–	2.49	2.59	Time:future
–	–	2.15	–	–	Time: present; simultaneous
2.11	–	–	–	–	Thought, belief
–	–	2.68	–	–	Trying hard
2.99	3.03	–	2.91	3.60	Using
–	2.87	–	–	–	Wanted
4.36	–	5.16	–	–	Worry

Table 5: Log ratio results across all statistically significant categories.