Measuring Privacy Concerns and the Right to Be Forgotten

Paul Steinbart
Arizona State University
paul.steinbart@asu.edu

Mark J. Keith
Brigham Young University
mark.keith@gmail.com

Jeffry S. Babb.
West Texas A&M University
jbabb@wtamu.edu

Abstract

The 'right to be forgotten' (RTBF) is an emerging concept that refers to an individual's ability to have data collected about themselves permanently deleted or "destroyed"—the final stage of the information life cycle. However, we do not yet understand where RTBF fits into existing theory and models of privacy concerns. This is due, at least in part, to the lack of validated instruments to assess RTBF. Therefore, following the methodology detailed by MacKenzie et al. [1], this paper develops scales to measure individuals' concerns about the RTBF. We validate the scale and show that the RTBF represents a separate dimension of privacy concerns that is not reflected in existing privacy concerns instruments.

1. Introduction

The explosive growth of 'Big Data' and the 'Internet of Things' means that ever more data about individuals is being collected, aggregated, and analyzed. Therefore it is not surprising that consumers have expressed an interest in being able to delete some of that information [2, 3]. For example, a recent national survey found that 88% of Americans supported a federal law mandating a right to delete their personal information that was collected and stored by organizations [2]. A similar poll in the EU reported that 75% of respondents wanted the ability to delete personal information [3]. In response, in 2014 the European Union Court of Justice ruled that Google had to provide some form of a RTBF to European consumers [4]. In the USA, California passed a law (SB 568) that provides a limited form of a RTBF to minors [5].

This desire for a RTBF is especially relevant for social networking sites, as evidence grows about the potential harm (e.g., loss of employment or education opportunities) resulting from information individuals posted about themselves [6, 7]. In response to this growing interest, Facebook allows users to delete their own profiles and search history [8]. Indeed, there is evidence that people are increasingly attempting to take control of their personal information that is posted on social networking sites [9], but are hampered in doing so because manual procedures are error-prone and tools are difficult to use [10]. Consequently, businesses have emerged (e.g., reputation.com) to perform such services.

However, attempting to monitor and defend one's online reputation is a reactive strategy that does not eliminate the potentially harmful information. Hence, consumer interest in a RTBF that would provide a more proactive strategy for managing one's online reputation. Indeed, the proximate impetus for establishing a RTBF in the EU stems from the desires of a Spanish citizen to restrict access to outdated information about the person's financial history [4]. The EU court acceded to those concerns by requiring Google to block that information from appearing in response to searches executed against the individual. However, because the information can still be accessed from the websites of the news organizations that originally published the stories, the court's ruling essentially provides for a right to be "de-indexed" (so that it does not appear in Google search results) rather than a pure right to delete information [11].

Although there has been much debate about the RTBF [6, 11-19], those discussions focus on questions of cost, feasibility, and likely impact on freedom of expression and the future development of the Internet. Scant attention has been paid to the relationship of the RTBF to consumers' privacy concerns. This is an important gap, given that the RTBF is intended to increase individuals' ability to protect their personal data.

Interestingly, existing instruments designed to measure privacy concerns [20-23] do not explicitly address the topic of data deletion. One explanation for this may be that these instruments were developed prior to the phenomenon of 'Big Data', when companies were still concerned about the cost of storage. Nevertheless, it is possible that the concerns that gave rise to the call for a RTBF are implicitly reflected in one or more of the existing scales. For example, the desire to be able to delete outdated information that is no longer relevant may be an aspect of wanting to control how one's personal information is used or part of concerns about being able to correct errors in that data. Both *control* and *errors* are aspects of the existing privacy concerns

URI: http://hdl.handle.net/10125/41765
ISBN: 978-0-9981331-0-2

scale [22, 24]. However, it is also possible that attitudes about the RTBF are not represented in existing measures of privacy concerns.

The question of whether the RTBF is already addressed in existing privacy concerns instruments or is a heretofore-neglected dimension of privacy is important because privacy concerns affect consumer intentions and behaviors [25-27]. Moreover, firms can choose to adopt a number of different attitudes toward the protection of consumer privacy [28] and can use privacy as a strategic competitive weapon that may enable them to charge higher prices [29]. Thus, if existing instruments omit an important dimension of privacy concerns, research results may be misleading and firms may make erroneous decisions. Consequently, our research questions are: *How should RTBF be measured?* and *How does the RTBF relate to previously identified dimensions of privacy concerns?*

To answer that question we follow MacKenzie et al.'s [1] prescriptions for construct development and develop and validate a scale to measure consumer attitudes about the RTBF. We then empirically test the relationship between RTBF and previously validated dimensions of privacy concerns.

2. Literature Review

Belanger and Crossler [30] and Smith et al. [31] reviewed and analyzed more than two decades of research on privacy. They note that one topic that has received considerable attention is the development of instruments to measure privacy concerns. However, although much progress has been made, an editorial that accompanied those two reviews identified the need to more precisely specify the nature of the construct *privacy concerns* [32, p. 984].

Subsequently, Hong and Thong [24] examined the questions used in prior research on privacy and validated a model in which individuals' privacy concerns consist of the following six dimensions:

1. Awareness that personal data is being collected and how it will be used
2. Collection of personal data
3. Control over the use of personal data
4. Secondary use and sharing of personal data with other entities
5. Protection of personal data from improper access
6. Errors in personal data and the ability to correct them

Collectively, the six dimensions address issues related to the acquisition, use, and storage of individuals' personal information, but say nothing about its disposal. This is surprising because the concept of the right to delete one's personal information has been discussed in legal journals [33], popular books [34], and privacy frameworks [35]. For example, Generally Accepted Privacy Principles [35] Principle 5 is titled "Use, Retention, and Disposal of Information" and recommends that "personal information is retained for no longer than necessary to fulfill the stated purposes [for which it was originally collected]" (section 5.2.2) and that "personal information no longer retained is anonymized, disposed of, or destroyed in a manner that prevents loss, theft, misuse, or unauthorized access" (section 5.2.3). Privacy advocates similarly stress the need for "end-to-end" protection of privacy throughout the entire information life cycle and the use of secure procedures to destroy personal information once it is no longer needed [36, 37].

Clearly, consumers are interested in some form of RTBF as a means to augment their ability to protect their personal data. Therefore, the question is whether those interests are reflected in the existing scales used to measure privacy concerns, or need to be added to those instruments. As previously mentioned, none of the six primary dimensions of privacy concerns explicitly mention deletion of data. Nevertheless, it is possible that one or more of the existing dimensions subsumes that issue.

The first two dimensions, awareness and collection, focus on the initial acquisition of personal information and, therefore, clearly do not address the issue of a RTBF. However, dimension three, control over how collected data is used, could possibly be interpreted as encompassing not just the processing of that data, but also its retention and disposal. Indeed, as previously noted, Generally Accepted Privacy Principle 5 addresses "use, retention, and disposal" of information. Nevertheless, examination of the specific questions used to measure control suggests otherwise. Two questions ask if it "usually bothers" the respondent when they "do not have control of personal information that I provide" and when they "do not have control or autonomy over decisions about how my personal information is collected, used, and shared", and the third question asks if the respondent is "concerned when control is lost or unwillingly reduced as a result of a marketing transaction" [24, p. 298]. Thus, it does not seem that any of the questions used to measure the "control" dimension of privacy concerns explicitly refer to the topic of the RTBF.

The fourth privacy concern dimension is secondary use. At first glance, this appears to at least partially address the issue of a RTBF, especially given that the case that triggered the EU's decision to create a RTBF revolved around the fact that outdated and, therefore, arguably irrelevant information about past financial conditions continued to appear in Google search results. However, the three questions used to measure concern about “secondary use” discuss concern about websites to which personal information is given using that information for unanticipated reasons, or selling it or sharing it with other entities [24, p. 297]. The EU case involving the Spanish citizen, however, dealt with information that was originally reported by news organizations, which then subsequently was indexed by Google not information provided by the data subject. Thus, it is not clear whether the existing questions about “secondary use” do address the RTBF.

Dimension five deals with concerns about improper access to stored personal information. Given the seemingly never-ending revelations about breaches that create the risk of identity theft, such concerns could indeed be a motivation for wanting to be able to delete information that entities store about oneself. However, although concern about protecting stored information may contribute to a desire for a RTBF in order to protect that data, it is not clear that the two constructs are identical.

Dimension six deals with concerns about errors in the data. The questions used to assess this dimension ask if the respondent is concerned that websites “do not take enough steps to make sure that my personal information in their files is accurate,” “do not have adequate procedures to correct errors in my personal information,” and “do not devote enough time and effort to verifying the accuracy of my personal information in their databases” [24, p. 298]. Those questions certainly are related to the concerns in the EU case that outdated and, therefore, irrelevant information about past financial history was returned in Google searches. However, as with concerns about secondary use and improper access, although concerns about errors may contribute to the demand for a RTBF, it is not clear that they are identical concepts.

Thus, it is not clear whether the six dimensions in existing instruments designed to measure privacy concerns address the issue of the RTBF. Therefore, we developed a scale designed to specifically address the RTBF and then empirically tested whether those questions load on one or more of the existing six dimensions of privacy concerns or represent a heretofore-overlooked dimension of privacy.

3. Methodology

We followed the methodology outlined by MacKenzie et al. [1] to develop a valid instrument (see Figure 1). In the results section we report the results for completing the first five phases (Steps 1-9 in Figure 1) of that process. In total, we conducted 4 separate data collections that were administered to various combinations of business school students and Amazon Mechanical Turk workers. Each subsection describes specific measures and populations for that phase of the project.

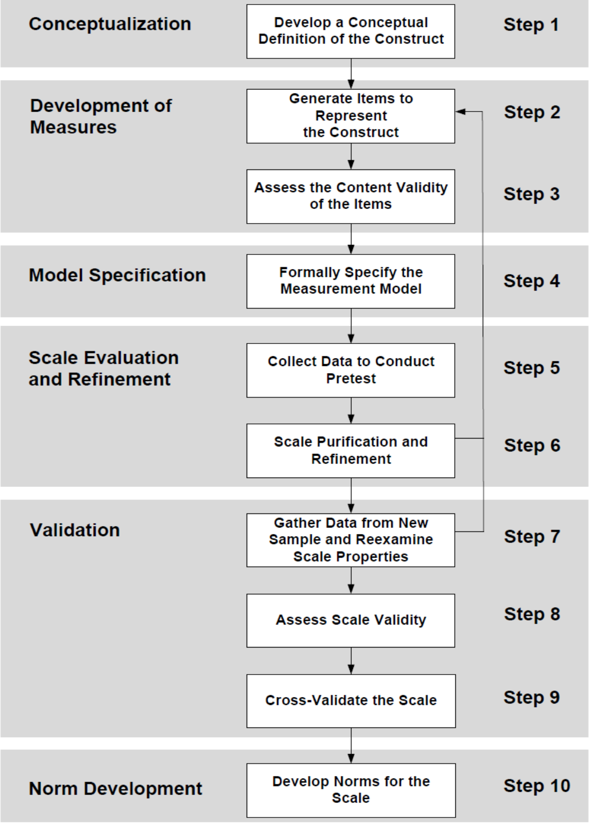
Conceptualization
Develop a Conceptual Definition of the Construct
Step 1
Development of Measures
Generate Items to Represent the Construct
Step 2
Assess the Content Validity of the Items
Step 3
Model Specification
Formally Specify the Measurement Model
Step 4
Scale Evaluation and Refinement
Collect Data to Conduct Pretest
Step 5
Scale Purification and Refinement
Step 6
Validation
Gather Data from New Sample and Reexamine Scale Properties
Step 7
Assess Scale Validity
Step 8
Cross-Validate the Scale
Step 9
Norm Development
Develop Norms for the Scale
Step 10

Figure 1. Instrument Development Process [1, p. 297]

4. Results

4.1. Conceptualization

Figure 1 shows that the first step in developing a new instrument entails developing a conceptual definition of the construct. We began by reviewing the discussions about the RTBF in the popular press and the academic literature. Based on our review of that literature we posit the following definition:

> *The RTBF applies to individuals. It is a feeling that reflects the desire to be able to delete personal information stored by other entities and accessible from the Internet.*

We suggest that the RTBF applies both to information that was directly provided by the individual [9, 17, 33, 34] and to information about an individual that was originally generated by others [4].

Prior research on privacy concerns has shown that those concerns vary in intensity depending upon the nature of the specific information in question (e.g., personal behaviors, beliefs, financial, health, etc.) [30, 31]. Discussions of the RTBF suggest that it, too, would vary depending upon the nature of the information in question. For example, the EU court case that requires Google to comply with the RTBF specifically involves financial information that is dated and, therefore, arguably irrelevant [4]. Similarly, discussions about the desire to delete social media posts focus on potentially embarrassing behaviors and statements of opinion [6, 9].

MacKenzie et al. [1, p. 302] note that constructs are not inherently formative or reflective, but that the way in which they are treated depends upon how the researcher conceptualizes the construct. The six previously validated dimensions of privacy concerns have been conceptualized as unidimensional constructs and, therefore, measured with a set of reflective indicators [24]. As noted earlier, those six dimensions address the acquisition, use and storage of personal information. The RTBF addresses concerns about the final stage of the information life cycle: disposal. Therefore, for purposes of testing whether the RTBF is already captured in the existing dimensions of privacy concerns, we treat it as a unidimensional construct and propose to measure it with a set of reflective indicators.

4.2. Development of Measures

Steps 2 and 3 entail creating items to measure the construct and assessing their content validity. Because the RTBF has not been previously measured, we created six initial items while drawing from the most recently validated instrument [24] for all other privacy concern items: *awareness* (AWA), *controls* (CON), *secondary use* (SEC), *errors* (ERR), *accuracy* (ACC), and *collection* (COL).

4.2.1 Content Adequacy Test

We then followed MacKenzie et al.'s [1, p. 304] recommended *content adequacy test* for assessing the content validity of the items by creating a matrix in which the columns represented different constructs and the rows represented items.

In our case, the matrix had seven columns, one for each of the previously validated dimensions of privacy concerns and one for the RTBF construct. The top row in each of those columns contained a definition of the construct. An eighth column, unlabeled, at the far left of the matrix contained individual question items in each row. The rows in the matrix contained the new items we created to measure the RTBF and the prior-validated items for the other six dimensions of privacy concerns.

We asked participants to rate, on a scale of 1-5, how well each item fit each dimension. Next, we ran a repeated measures analysis of variance (rANOVA) to analyze each individual survey item to see if it rates significantly highest on its own sub-construct by examining a contrast comparison between the intended sub-construct and all others. If the difference is statistically significant, then the content is valid and the item is more likely to demonstrated discriminant validity [38].

Two content adequacy tests were performed on our six new items as well as the existing 18 items from the most recently validated privacy concerns scale [24]. To our knowledge, this is the first time that the privacy concerns scale has been analyzed using the content adequacy test. Because of the high cognitive load of rating each item across each sub-construct (7 x 24 = 168 questions), we made three versions of the survey that each included a random selection of 8 of the 24 items to rate across the sub-constructs. The survey also included the actual privacy concerns instrument after the content adequacy test for a total of 80 items plus demographic questions. The results were then combined into a single data set. The results of the first data collection included 800 responses from university students in the business college of a large public university in the western United States. However, 231 responses were removed for being incomplete, incorrectly answering trap questions, straight-lining, or taking very little time to complete the entire survey, resulting in 569 usable responses.

4.2.2. Content Adequacy Test Results

Table 1 summarizes the results, which were mostly positive with some exceptions. The gray shaded cells indicate an item loading significantly higher on its own intended construct than all others. The black cells with white text indicate problems with an item. First, RTBF3 did not load significantly higher on the RTBF construct than on CON. Therefore, it was removed from subsequent data collections. However, the five remaining items were valid. Interestingly, one of the existing scale items for the sub-construct AWA and another for COL did not pass the content adequacy test even though they were copied directly from past research [22, 24]. AWA3 did load highest on its own factor, but that loading was not significant in the rANOVA. COL3 did not even load highest on its own sub-construct.

Table 1. Results of First Content Adequacy Test

	RTBF	ACC	AWA	COL	CON	ERR	SEC
RTBF1	3.97	3.11	3.24	3.17	3.77	2.78	2.89
RTBF2	3.98	3.03	3.33	3.16	3.78	2.81	2.94
RTBF3	3.89	2.91	3.36	3.05	3.83	2.70	2.89
RTBF4	3.94	3.04	3.28	3.08	3.68	2.76	2.88
RTBF5	4.15	2.98	3.50	3.18	3.66	2.86	2.93
RTBF6	4.07	2.95	3.56	3.14	3.76	2.87	2.82
ACC1	3.01	3.98	3.30	3.01	3.29	2.68	3.32
ACC2	3.05	4.04	3.32	2.92	3.23	2.49	3.14
ACC3	3.14	3.99	3.60	3.22	3.46	2.76	3.51
AWA1	3.16	3.29	3.76	3.15	3.38	2.81	3.24
AWA2	3.02	3.34	4.03	3.23	3.63	2.68	3.52
AWA3	3.04	3.36	3.67	3.63	3.64	2.80	3.23
COL1	3.13	3.29	3.34	3.90	3.36	2.61	3.08
COL2	2.96	2.87	3.16	3.63	3.25	2.44	3.02
COL3	3.32	3.20	3.70	3.63	3.54	2.84	3.22
CON1	3.38	3.24	3.54	3.40	4.11	2.83	3.20
CON2	3.23	3.31	3.60	3.57	4.10	2.68	3.60
CON3	3.07	3.35	3.50	3.16	3.82	2.88	3.24
ERR1	2.94	3.28	3.07	3.35	2.80	3.82	2.77
ERR2	2.75	3.14	2.91	3.22	2.77	3.87	2.64
ERR3	2.76	3.29	3.18	3.27	2.81	3.85	2.72
SEC1	3.49	3.37	3.45	2.67	3.54	3.17	3.94
SEC2	3.42	3.21	3.39	2.50	3.42	3.08	3.97
SEC3	3.59	3.34	3.61	2.70	3.80	3.24	3.93

Therefore, to further contribute to the body of research on the privacy ,s instrument, we made slight modifications to those existing scales (the entire scale is found later) and executed another content adequacy test. In the second round of testing, we collected 269 responses (238 usable) from Amazon Mechanical Turk and only included AWA3, COL3, and 6 other randomly selected items to keep the survey length comparable to the prior data collection. The ANOVAs revealed that our changes to AWA3 and COL3 improved them enough to rate significantly highest on their own intended sub-constructs (p < 0.001). Therefore, we conclude that our scale (including the five remaining RTBF items created for this study) has sufficient content validity.

4.3. Model Specification

Step 4 in the instrument creation process involves specifying the formal measurement model. In the most recent privacy concerns instrument validation, Hong and Thong [24] found evidence of a third order privacy concern factor with two second-order factors among the 6 prior sub-constructs. Figure 1 visualizes the model specification they selected after testing 12 alternatives.

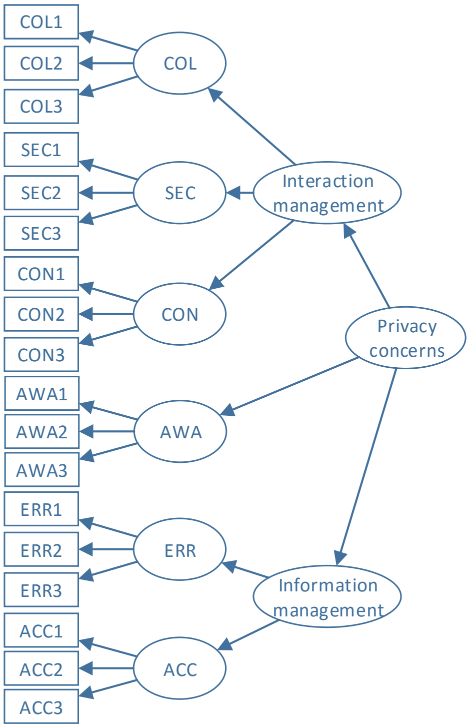

Figure 1. Privacy Concern Model [24]

An important aspect of this research is to discover where RTBF fits within this existing privacy concern model. Therefore, after the RTBF scale is validated, we will examine whether it fits best as: 1) a reflection directly from the third order privacy concern factor, 2) a reflection of the second order *interaction management* factor, 3) a reflection of the second order *information management* factor, or 4) an entirely separate factor from privacy concern.

4.4. Scale Evaluation and Refinement

Steps 5 and 6 of the instrument creation process entail collecting data to pretest the instrument and then using those results to purify and refine the instrument, respectively.

Before the two content adequacy tests described previously, we executed a pilot test which led to several important changes even before the content adequacy tests began. In addition, the entire privacy concerns scale was measured during each content adequacy test for a total of three unique pilot tests.

With each pilot test, we performed a covariance-based structural equation modeling tool. Table 2 summarizes the samples and CFA results of each pilot test. The results indicate adequate fit [1, 39].

Table 2. CFA results for each pilot tests

	Pilot 1	Pilot 2	Pilot 3
n	224	569	238
Source	AMT	students	AMT
CMIN/df	2.082	1.918	2.002
NFI	.860	.924	.875
CFI	.964	.962	.933
PRatio	.826	.757	.826
RMSEA	.070	.042	.071

The following statistics were calculated solely for the last pilot test. Reliability for each scale was analyzed by measuring Cronbach's alpha for each of the sub-dimensions and was well-above the 0.7 threshold, ranging from 0.87 to 0.93.

To evaluate the reliability of each individual scale item, we examined the significance of the estimate (λ) of the relationship between an indicator and the latent construct. All items were significant.

Convergent validity was analyzed by calculating the average variance extracted (AVE) for each first-order reflective sub-construct. All AVEs were well-over the 0.50 recommended cutoff [40], ranging from 0.76 to 0.86. The second order privacy s construct was also above the cutoff at 0.78. Convergent validity for this data was sufficient.

Discriminant validity was analyzed by examining whether the average variance explained (AVE) by the indicators for their underlying latent constructs is greater than the squared correlation between the focal construct and the other sub-constructs [40]. The results indicated sufficient discriminant validity as all AVEs for each sub-construct were greater than their squared correlations with other sub-constructs.

In summary, we conclude that the pilot data—after accounting for the changes made during the content adequacy tests—exhibits sufficient reliability to proceed with final data collection.

4.5. Validation

Validation of the refined scale (steps 7 through 9 in the instrument creation process) includes the collection of new data, validation of the scale with the new data, and cross-validation from different populations [1].

The final data collection was based on a combined sample of 331 AMT workers and 78 students. As with the pilot data collections, responses sets were removed if participants: 1) straight-lined responses, 2) missed any of the four trap questions, or 3) spent less than 1/3 of the median time taking the survey. As usual, all latent construct items were completely randomized across all constructs and sub-constructs. This resulted in a total of 324 completed response sets. Demographic questions concerning age, current residence, ethnicity, and education were included at the end of the survey.

Because no changes were made to the scales after the third pilot test, the data were combined for a total of 552 response sets. The participants were 57 percent male and an average of 35 years old. Six percent were drawn from people currently living outside of the US. 77 percent of respondents were Caucasian, 6 percent African American, 3 percent Hispanic, and 11 percent Asian. On average, participants had earned at least a 2-year degree.

4.5.1. Final Measurement Model

The scale validity statistics were slightly improved over the pilot tests. The CFA produced the following results: CMIN/df = 2.733, NFI = 0.951, CFI = 0.968, PRatio = 0.826, and RMSEA = 0.056.

Reliability for each sub-construct was high with Cronbach's alpha ranging from 0.84 to 0.92. All scale item estimates (λ) were significant. All AVEs exceeded 0.50 [40], ranging from 0.75 to 0.84. Also, all AVEs exceeded the squared correlation of each sub-construct with every other sub-construct. In summary, the final data set indicated strong reliability and validity. Table 3 lists the final measurement scale including both the RTBF items. It should also be noted that the items AWA3 and COL3, while drawn from prior research [24], were updated in this study based on the results of the content adequacy tests (see Table 1).

Table 3. Final Privacy Concerns Scale Including RTBF

ACC1: I am concerned that company or government agencies do not protect my personal information from unauthorized access.
ACC2: I am concerned that companies or government agencies do not devote enough time and effort to preventing unauthorized access to my personal information.
ACC3: I am concerned that companies or government agencies do not take enough steps to make sure that unauthorized people cannot access my personal information in their computers.
AWA1: I am concerned when a clear and conspicuous disclosure is not included in the privacy policies of companies or government agencies.
AWA2: It usually bothers me when I am not aware or knowledgeable about how my personal information will be used by companies or government agencies.
AWA3: It usually bothers me when companies or government agencies do not tell me the way the data are collected, processed, and used.
COL1: I am concerned that companies or government agencies are collecting too much personal information about

me.
COL2: It usually bothers me when companies or government agencies ask me for personal information.
COL3: When companies or government agencies try to collect my personal information, I sometimes hesitate to provide it.
CON1: It usually bothers me when I do not have control of personal information that I provide to companies or government agencies.
CON2: It usually bothers me when I do not have control or autonomy over decisions about how my personal information is collected, used, and shared by companies or government agencies.
CON3: I am concerned when control is lost or unwillingly reduced as a result of a marketing transaction with companies or government agencies.
ERR1: I am concerned that companies or government agencies do not take enough steps to make sure that my personal information in their files is accurate.
ERR2: I am concerned that companies or government agencies do not have adequate procedures to correct errors in my personal information.
ERR3: I am concerned that companies or government agencies do not devote enough time and effort to verifying the accuracy of my personal information in their databases.
RTBF1: I am concerned that companies or government agencies do not allow me to delete information I've given them.
RTBF2: It usually bothers me that companies or government agencies don't offer a process for me to request deletion of information I've given them.
RTBF4*: I am concerned that companies or government agencies may not honor my requests to delete information I've given them.
RTBF5*: It usually bothers me that companies or government agencies do not give me the option to have my information deleted.
RTBF6*: I am concerned that companies or government agencies may not be capable of deleting my information when I request that they do so.
SEC1: I am concerned that when I give personal information to a company of government agency, the entity would use the information for other reasons.
SEC2: I am concerned that companies or government agencies would sell my personal information to other companies.
SEC3: I am concerned that companies or government agencies would share my personal information with other entities without my consent.

Note: *RTBF3 was removed based on the content adequacy test. All remaining RTBF items are valid. However, to reduce survey fatigue, we recommend retaining RTBF4, RTBF5, and RTBF6 which exhibited the best discriminant validity. Bolded terms (AWA3 and COL3) represent modified versions of previously validated questions that improved the performance of the content adequacy test.

4.5.2. Determining Model Specification

After confirming the measurement scale validity and reliability statistics for the new RTBF scale (along with the other six privacy concern sub-constructs), we next seek to determine the most appropriate model specification. In particular, we next examine model fit statistics with four alternative models (depicted in Figure 2) of how the RTBF could relate to the previously validated structure of privacy concerns [24].

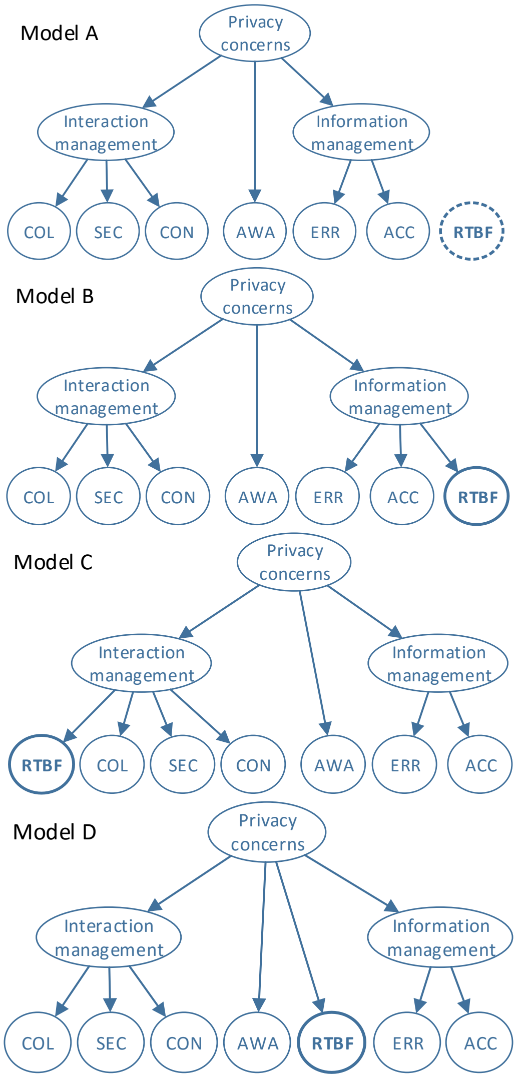

Figure 2. Alternative Model Specifications

Model 1 treats the RTBF as entirely separate from privacy concerns. Model 2 includes RTBF as a sub-construct of the second order factor *interaction management*. Model 3 includes RTBF as a sub-construct of the second order factor *interaction* *management.* Model 4 includes RTBF as a reflection directly from the third order *privacy concerns* factor. Table 4 summarizes the model fit statistics for each version of the models in Figure 2.

Table 4. Fit Statistics across Models

	1	2	3	4
CMIN/df	2.882	2.932	2.909	2.914
NFI	0.958	0.945	0.945	0.945
CFI	0.972	0.963	0.963	0.963
PRATIO	0.743	0.801	0.801	0.801
RMSEA	0.058	0.059	0.058	0.058

Based on the fit statistics in Table 4, each model of privacy concerns demonstrates good fit. Therefore, it appears that the RTBF is indeed related to the other six dimensions of privacy concerns, but the exact nature of that relationship needs further investigation.

5. Discussion and Conclusion

Prior research on privacy has established that people are concerned about being adequately aware of the information being collected about themselves, the amount and nature of the information that is collected, the degree to which they can control what organizations do with their information, secondary use and sharing of that information, the ability to correct errors in information stored about them, and the security of that information. This study investigated whether those six dimensions of privacy concerns also subsume the recent interest in the ability to delete personal information, referred to as the RTBF, or if the RTBF is a separate construct.

We followed recommended procedures MacKenzie et al. [1] to develop and validate a set of five questions to assess an individual's feelings about the RTBF. We then included those five questions along with previously validated questions about the six existing dimensions of privacy concerns and administered the instrument to both AMT workers and students. Our results show that the RTBF is a separate construct that is distinct from the existing six dimensions of privacy concerns.

Thus, our results indicate that existing measures of privacy concerns do not incorporate people's concerns about the RTBF. Therefore, research about the antecedents and outcomes of privacy concerns may not necessarily apply to the RTBF. Consequently, organizations should not assume that their existing privacy strategy [e.g., 41] adequately addresses consumers' feelings about the RTBF. A reasonable explanation for our findings regarding RTBF is that the phenomenon of interest is emergent and not fully calibrated into the experiences of everyday life. Similarly, it was not long ago that being in possession of a social media profile was a non-topic to begin with.

Nevertheless, the RTBF is highly correlated with existing dimensions of privacy concerns. One likely explanation is that all seven constructs deal with an individual's personal information that is stored with various organizations. However, the RTBF appears to be distinct from the six dimensions of privacy concerns identified by prior research. This may reflect a critical difference in the temporal relevance associated with those seven constructs. Some of the other six dimensions of privacy concerns represent factors that have immediate bearing on decisions about whether to disclose or share information: To what extent do I feel adequately aware that information about me is being collected (AWA)? How comfortable am I in disclosing specific types of information (COL)? How secure will that information be (ACC)? Other questions have delayed, but still relatively short-term relevance: How might my information be subsequently shared (SEC)? Will I be able to correct errors (ERR)? How much control will I be able to exert over what the organization does with my information (CON)?

In contrast, the RTBF is likely to become relevant only at some time in the future, when someone is contemplating changes in behavior or status. This is important because psychology research shows that people tend to focus primarily on foreseeable and imaginable costs and benefits when making decisions, and ignore or grossly underweight factors that are not relevant until much later [42, 43]. For example, although people may be able to weigh, with varying degrees of accuracy, the costs and benefits of adopting a health-related behavior, it is much more difficult to accurately consider the difficulty of changing that behavior at some unspecified time in the future [43]. Our results suggest that it may be the same with decisions about the disclosure of personal information: the immediate and short-term costs and benefits are evaluated separately, yet related to, the possibility of wanting to delete that information at some later time. If so, just as research on the use of IT has had to expand beyond the study of initial adoption and explicitly address issues of discontinuance [44, 45], privacy research must explicitly address not only decisions about whether to disclose personal information, but also explicitly investigate what prompts people to desire to discontinue such disclosure.

As with any research, it is important to acknowledge this study's limitations. Some may criticize a reliance on students and AMT respondents in our pilot studies. As this work progresses forward, some redress of this will be undertaken with a more robust and diverse sampling. However, recent research has indicated that AMT participants classified as “master workers” (which was the case in this study) are at least as valid as those collected from professional data collection services [46, 47].

Additionally, we have not yet tested the nomological relationship between RTBF, privacy concerns, and other factors. We have also not resolved the question of how the RTBF relates to the other six dimensions of privacy concerns: all four models we tested fit the data well. Furthermore, we have not yet examined how attitudes about the RTBF differ across types of information. Consumers have different levels of privacy concerns for different types of information [48, 49]. Consequently, it is not surprising that there is some evidence that European consumers’ requests to apply the EU’s recently established RTBF to have Google delist search engine results focus on some types of information more than others [50]. However, there are many other relevant forms of information that consumers may want deleted besides that which Google indexes (e.g., private social media posts restricted to friends, Internet of things data, mobile device sensory data, etc.). Therefore, additional research is needed to more fully understand how consumers’ feelings about the RTBF vary across types of information besides that which shows up in search engines. Such research may help resolve the ongoing heated debates about whether the RTBF conflicts with first amendment rights (c.f. [14] versus [51]) because they want to censor news stories about themselves or whether they are more interested in being able to control retention and storage of personal information they post on social media and that businesses collect about their online behaviors. We plan to address all these issues in subsequent work.

In conclusion, this study developed and validated a set of questions that can be used to assess a person’s feelings about the RTBF. Those feelings are related to, but distinct from, previously validated dimensions of privacy concerns.

6. References

[1] Mackenzie, S.B., Podsakoff, P.M., and Podsakoff, N.P., "Construct Measurement and Validation Procedures in Mis and Behavioral Research: Integrating New and Existing Techniques", Mis Quarterly, 35(2), 2011, pp. 293-334.

[2] Allstate, "New Poll Shows Americans Anxious About Privacy", in (Editor, 'ed.'^'eds.'): Book New Poll Shows Americans Anxious About Privacy, 2013

[3] Commission, E., "Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union", in (Editor, 'ed.'^'eds.'): Book Special Eurobarometer 359: Attitudes on Data Protection and Electronic Identity in the European Union, 2011

[4] Commission, E., "Factsheet on the ‘Right to Be Forgotten’ Ruling”, C-135/12, European Commission, 2015", in (Editor, 'ed.'^'eds.'): Book Factsheet on the ‘Right to Be Forgotten’ Ruling”, C-135/12, European Commission, 2015, 2015

[5] Musil, S., "California Gives Teens an ‘Eraser Button’ to Hide Online Skeletons", in (Editor, 'ed.'^'eds.'): Book California Gives Teens an ‘Eraser Button’ to Hide Online Skeletons, 2013

[6] Rosen, J., "The Web Means the End of Forgetting", in (Editor, 'ed.'^'eds.'): Book The Web Means the End of Forgetting, 2010

[7] Bradley, T., "What You Don't Know About Your Online Reputation Can Hurt You", in (Editor, 'ed.'^'eds.'): Book What You Don't Know About Your Online Reputation Can Hurt You, 2010

[8] Constine, J., "Facebook Starts Letting You View and Delete Your Facebook Search History", in (Editor, 'ed.'^'eds.'): Book Facebook Starts Letting You View and Delete Your Facebook Search History, 2012

[9] Madden, M., "Privacy Management on Social Media Sites", in (Editor, 'ed.'^'eds.'): Book Privacy Management on Social Media Sites, 2012

[10] Yang, S., "Understanding the Pain: Examining Individuals' Online Reputation Management Behaviour and Its Obstacles--a Grounded Theory", in (Editor, 'ed.'^'eds.'): Book Understanding the Pain: Examining Individuals' Online Reputation Management Behaviour and Its Obstacles--a Grounded Theory, IEEE, 2016, pp. 3898-3907.

[11] Bygrave, L.A., "A Right to Be Forgotten?", Communications of the ACM, 58(1), 2015, pp. 35-37.

[12] Benett, S.C., "The Right to Be Forgotten: Reconciling Eu and Us Perspectives", Berkeley Journal of International Law, 30(2012, pp. 1.

[13] De Hert, P., and Papakonstantinou, V., "The Proposed Data Protection Regulation Replacing Directive 95/46/Ec: A Sound System for the Protection of Individuals", Computer Law & Security Review, 28(2), 2012, pp. 130-142.

[14] Larson Iii, R.G., "Forgetting the First Amendment: How Obscurity-Based Privacy and a Right to Be Forgotten Are Incompatible with Free Speech", Communication Law and Policy, 18(1), 2013, pp. 91-120.

[15] Mantelero, A., "The Eu Proposal for a General Data Protection Regulation and the Roots of the ‘Right to Be Forgotten’", Computer Law & Security Review, 29(3), 2013, pp. 229-235.

[16] Newman, A.L., "What the “Right to Be Forgotten” Means for Privacy in a Digital Age", Science, 347(6221), 2015, pp. 507-508.

[17] Rosen, J., "The Right to Be Forgotten", Stanford law review online, 64(2012, pp. 88.

[18] Rustad, M.L., and Kulevska, S., "Reconceptualizing the Right to Be Forgotten to Enable Transatlantic Data Flow", Harvard Journal of Law and Technology, 28(2015, pp. 349.

[19] Weber, R., "The Right to Be Forgotten: More Than a Pandora’s Box? In 2 Jipitec 120, 121", in (Editor,

'ed.'^'eds.'): Book The Right to Be Forgotten: More Than a Pandora’s Box? In 2 Jipitec 120, 121, 2011

[20] Dinev, T., and Hart, P., "Internet Privacy Concerns and Their Antecedents-Measurement Validity and a Regression Model", Behaviour & Information Technology, 23(6), 2004, pp. 413-422.

[21] Malhotra, N.K., Kim, S.S., and Agarwal, J., "Internet Users’ Information Privacy Concerns (Iuipc): The Construct, the Scale, and a Causal Model", Information Systems Research, 15(4), 2004, pp. 336-355.

[22] Smith, H.J., Milberg, S.J., and Burke, S.J., "Information Privacy: Measuring Individual's Concerns About Organizational Practices", Mis Quarterly, 20(2), 1996, pp. 167-196.

[23] Stewart, K.A., and Segars, A.H., "An Empirical Examination of the Concern for Information Privacy Instrument", Information Systems Research, 13(1), 2002, pp. 36-49.

[24] Hong, W., and Thong, J.Y., "Internet Privacy Concerns: An Integrated Conceptualization and Four Empirical Studies", Mis Quarterly, 37(1), 2013, pp. 275-298.

[25] Angst, C.M., and Agarwal, R., "Adoption of Electronic Health Records in the Presence of Privacy Concerns: The Elaboration Likelihood Model and Individual Persuasion", Mis Quarterly, 33(2), 2009, pp. 339-370.

[26] Dinev, T., and Hart, P., "An Extended Privacy Calculus Model for E-Commerce Transactions", Information Systems Research, 17(1), 2006, pp. 61-80.

[27] Hui, K.L., Teo, H.H., and Lee, S.Y.T., "The Value of Privacy Assurance: An Exploratory Field Experiment", Mis Quarterly, 31(1), 2007, pp. 19-33.

[28] Greenaway, K.E., Chan, Y.E., and Crossler, R.E., "Company Information Privacy Orientation: A Conceptual Framework", Information Systems Journal, 25(6), 2015, pp. 579-606.

[29] Mai, B., Menon, N.M., and Sarkar, S., "No Free Lunch: Price Premium for Privacy Seal-Bearing Vendors", Journal of Management Information Systems, 27(2), 2010, pp. 189-212.

[30] Belanger, F., and Crossler, R.E., "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems", Mis Quarterly, 35(4), 2011, pp. 1017-1041.

[31] Smith, H.J., Dinev, T., and Xu, H., "Information Privacy Research: An Interdisciplinary Review", Mis Quarterly, 35(4), 2011, pp. 989-1015.

[32] Pavlou, P.A., "State of the Information Privacy Literature: Where Are We Now and Where Should We Go?", Mis Quarterly, 35(4), 2011, pp. 977-988.

[33] Blanchette, J.-F., and Johnson, D.G., "Data Retention and the Panoptic Society: The Social Benefits of Forgetfulness", The Information Society, 18(1), 2002, pp. 33-45.

[34] Mayer-Schönberger, V., Delete: The Virtue of Forgetting in the Digital Age, Princeton University Press, 2011.

[35] Aicpa/Cica, Generally Accepted Privacy Principles, American Institute of Certified Public Accountants, New York, 2009.

[36] Cavoukian, A., "Operationalizing Privacy by Design: A Guide to Implementing Strong Privacy Practices", Information and Privacy Commissioner, Ontario, Canada, 2012,

[37] Cavoukian, A., "Privacy by Design and the Emerging Personal Data Ecosystem", Privacy By Design, 2012,

[38] Hinkin, T.R., and Tracey, J.B., "An Analysis of Variance Approach to Content Validation", Organizational Research Methods, 2(2), 1999, pp. 175-186.

[39] Anderson, J.C., and Gerbing, D.W., "Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach", Psychological bulletin, 103(3), 1988, pp. 411-423.

[40] Fornell, C., and Larcker, D.F., "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error", Journal of Marketing Research, 18(1981), 1981, pp. 39-50.

[41] Lee, D.-J., Ahn, J.-H., and Bang, Y., "Managing Consumer Privacy Concerns in Personalization: A Strategic Analysis of Privacy Protection", Mis Quarterly, 35(2), 2011, pp. 423-444.

[42] Loewenstein, G., and Prelec, D., "Anomalies in Intertemporal Choice: Evidence and an Interpretation", Quarterly Journal of Economics, 107(2), 1992, pp. 573-597.

[43] Mischel, W., Ebbesen, E.B., and Raskoff Zeiss, A., "Cognitive and Attentional Mechanisms in Delay of Gratification", Journal of personality and social psychology, 21(2), 1972, pp. 204.

[44] Steinbart, P.J., Keith, M.J., and Babb, J., "Examining the Continuance of Secure Behavior: A Longitudinal Field Study of Mobile Device Authentication", Information Systems Research, forthcoming(2016,

[45] Bhattacherjee, A., "Understanding Information Systems Continuance: An Expectation-Confirmation Model", Mis Quarterly, 25(3), 2001, pp. 351-370.

[46] Buhrmester, M., Kwang, T., and Gosling, S.D., "Amazon's Mechanical Turk a New Source of Inexpensive, yet High-Quality, Data?", Perspectives on psychological science, 6(1), 2011, pp. 3-5.

[47] Peer, E., Vosgerau, J., and Acquisti, A., "Reputation as a Sufficient Condition for Data Quality on Amazon Mechanical Turk", Behavior Research Methods, 46(4), 2014, pp. 1023-1031.

[48] Milne, G.R., Pettinico, G., Hajjat, F.M., and Markos, E., "Information Sensitivity Typology: Mapping the Degree and Type of Risk Consumers Perceive in Personal Data Sharing", Journal of Consumer Affairs, 2016,

[49] Sheehan, K.B., and Hoy, M.G., "Dimensions of Privacy Concern among Online Consumers", Journal of public policy & marketing, 19(1), 2000, pp. 62-73.

[50] Xue, M., Magno, G., Cunha, E., Almeida, V., and Ross, K.W., "The Right to Be Forgotten in the Media: A Data-Driven Study", Proceedings on Privacy Enhancing Technologies, 2016(4), 2016, pp. 389-402.

[51] Burkell, J.A., "Remembering Me: Big Data, Individual Identity, and the Psychological Necessity of Forgetting", Ethics and Information Technology, 18(1), 2016, pp. 17-23.