Taylor & Francis
Taylor & Francis Group

Journal of Computer Information Systems

ISSN: 0887-4417 (Print) 2380-2057 (Online) Journal homepage: http://www.tandfonline.com/loi/ucis20

Explaining Diversity and Conflicts in Privacy Behavior Models

Anna Rohunen, Jouni Markkula, Marikka Heikkilä & Markku Oivo

To cite this article: Anna Rohunen, Jouni Markkula, Marikka Heikkilä & Markku Oivo (2018): Explaining Diversity and Conflicts in Privacy Behavior Models, Journal of Computer Information Systems

To link to this article: https://doi.org/10.1080/08874417.2018.1496804

Published online: 01 Aug 2018.

JOURNAL OF COMPUTER INFORMATION SYSTEMS
https://doi.org/10.1080/08874417.2018.1496804

Taylor & Francis
Taylor & Francis Group

Explaining Diversity and Conflicts in Privacy Behavior Models

Anna Rohunen[a], Jouni Markkula[a], Marikka Heikkilä[b], and Markku Oivo[a]

[a]University of Oulu, Oulu, Finland; [b]University of Turku, Turku, Finland

ABSTRACT

Technological development and increasing personal data collection and utilization raise the importance of understanding individuals' privacy behavior. Privacy behavior denotes the willingness to disclose personal data for services utilizing these data. The literature presents various privacy behavior models (PBMs). However, the research is incoherent, with inconsistencies among models. Therefore, the application and subsequent development of PBMs are challenging. Different background theories are used for model construction, and studies have been conducted in distinct application domains. We studied whether the models' inconsistencies could be explained by these differences. Our in-depth analysis of PBMs was based on a systematic literature review of the most often cited key studies. Our findings indicate that the choice of theories and the application domains do not explain inconsistencies; instead, the models are often of an ad hoc type and constructed in an eclectic way. These results imply the need for more consistent research on privacy behavior.

KEYWORDS

Privacy behavior; privacy behavior models; empirical models; personal data intensive services

Introduction

Information privacy has become an increasingly important research topic, along with technological development and evolving personal data collection contexts, such as intelligent traffic systems or electronic services. In these contexts, individuals' data are collected for commercial transactions and the production of personal data intensive services and are often also used for marketing purposes. Individuals' privacy behavior denotes their adoption and usage of services and technologies that require personal data disclosure, as well as indicates their willingness to disclose their data for these services. The literature has presented both theoretical and empirical models that investigate privacy behavior antecedents (i.e., variables that affect privacy behavior) and explain their relationships to data subjects' behavioral outcomes, such as their willingness to disclose data. The existing research comprises an extensive body of empirical, quantitative models that aim to describe and explain privacy behavior (cf. refs. 1 and 2). We refer to these as privacy behavior models (PBMs), which identify a broad range of privacy behavior antecedents. PBMs rest on the data subjects' view of their privacy. Information privacy can also be perceived from other aspects, such as social or economic (i.e., utilization of personal information may have implications for not only individual persons but also different organizations or the whole society). Due to our research topic, we focus on the individuals' view. Information privacy concerns associated with personal data collection and usage are among the PBMs' key antecedents affecting privacy behavior by decreasing the willingness to use services that require personal data disclosure. Privacy behavior antecedents also include data subjects' experienced data disclosure risk and their trust in data collecting organizations (cf. refs. 3–7). The models often consider economic or social benefits of the disclosure as well, such as improved service personalization and the common good supported by the data.[8,9] Some models have focused on privacy behavior antecedents different from these, such as sensitivity of the collected data and data subjects' personality traits.[10–12]

The PBM research has been relatively fragmented (cf. refs. 1, 2, and 13). Existing PBMs sometimes conflict with one another regarding the relationships among their privacy behavior antecedents. In recent years, some literature reviews have been conducted to synthesize the PBMs [1,13]. However, these reviews have been descriptive and have not investigated the reasons for the differences among the models. For both privacy researchers and practitioners, it is highly relevant to understand the differences. In research, this understanding is needed in PBM construction and development to create usable models with appropriate contextualization and improved validity in explaining privacy behavior. From the practical perspective, understanding the models helps in interpreting and applying them and guides practitioners in choosing a suitable PBM for a particular situation. As different types of background theories are currently used for PBM construction and individual studies have also been conducted in various application domains, the question is raised about whether these could explain the models' differences.

To gain a better understanding of the differences among the PBMs, we studied whether they could be explained by the different background theories and application domains used for designing the models. To answer this question, we conducted a literature review and an in-depth analysis of

CONTACT Anna Rohunen ✉ anna.rohunen@oulu.fi University of Oulu, Faculty of Information Technology and Electrical Engineering (ITEE), Empirical Software Engineering in Software, Systems and Services (M3S) research unit, P.O. Box 4500, FI-90014 Finland.

the most cited key PBM studies. Through our review and analysis, we aimed to identify background theories and analyze their usage in PBM development, as well as investigate how the data collection domain had been taken into account in the existing PBMs.

The rest of the paper is organized as follows. The related research section presents an overview of the existing PBM research. The research methods of our literature review and analysis are explained in the research methodology section. The results section summarizes both the PBMs identified in the literature review and their background theories, as well as describes application of theories and how the application domains have been considered in the PBMs. This study's findings are then discussed with respect to both their theoretical and practical implications. The last section concludes the paper.

Related research

The PBM research comprises a wide variety of empirical models, constructed across several disciplines (cf. refs. 1 and 2) These models have substantial differences in their incorporated privacy behavior antecedents, as well as in the roles and the relationships among these antecedents. Inconsistencies exist among these PBMs, with some even conflicting with one another in explaining privacy behavior. For example, antecedents may affect privacy behavior, either in parallel or mediate one another in the PBMs, and their relationships may be controversial among separate models (see, e.g., refs. 4–8). On these grounds, PBM research seems fragmented, and overall, it lacks replication and synthesis (cf. refs. 1, 2, and 13).

In recent years, there has been a tendency to conduct integrative literature reviews on information privacy behavior antecedents and correspondingly, behavioral outcomes in relation to information disclosure. These literature reviews have aimed to synthesize the existing research. Based on them, macro models centering on information privacy concerns have been developed to describe the relationships among the PBM constructs.[1,2,14] However, these macro models also differ in their constructs and in the constructs' conceptualizations. Based on Li's review of the theories used in online information privacy research,[13] he developed an integrated framework that summarized the interrelationships among 15 established theories and outlined the relationships among privacy behavior variables. However, research of this type alone does not have enough potential to explain differences and conflicts among existing models.

Fragmentary knowledge in PBM research makes subsequent model construction and development more difficult. It also impedes interpretation and application of existing models, making their utilization for practical purposes troublesome. Since the earlier literature reviews only aimed to synthesize the existing PBMs, an in-depth investigation is needed to explain the reasons behind the conflicting models. As the distinct PBMs apply diverse background theories and focus on various data collection domains, these differences could potentially explain the inconsistencies among the models. For this reason, the roles of background theories and different types of data collection domains in PBM design should be examined in more detail. We assume that the models that have been designed based on a common solid theoretical background, rigorous application of research methods, and appropriate contextualization to a specific domain can be expected to be mature and systematically of higher quality than various *ad hoc*-type models.

In this study, we aimed to respond to the need to explain model inconsistencies through a review and an analysis of the existing PBMs. We systematically reviewed the journal articles published until the end of 2017, as well as identified and analyzed the most valid and high-quality PBMs with respect to their background theories and personal data collection domains. Our paper contributes to PBM construction and development by providing insights into the application of the theories and how to take into account different types of data collection domains in privacy behavior modeling.

Research methodology

We aimed to explore the reasons for inconsistencies and conflicts among existing PBMs and provide initial insights into these differences in a formal way. For this purpose, we intended to find and analyze the most cited key PBM studies with solid theoretical backgrounds and rigorous application of research methods. Our objective was first to identify a focused set of the most valid and high-quality key papers and then delve deeper into their theory usage and application domains. An in-depth analysis of this type, with an extensive set of papers, would have been unfeasible due to its demand for substantial effort and resources. On the other hand, analyses based on a relatively small number of studies might bring on validity issues and lead to selection bias. To mitigate the possibility of this bias, we obtained our set of key papers by applying systematic literature review (SLR) principles and techniques following Kitchenham and Charters' guidelines.[15]

We adapted Kitchenham and Charters' guidelines[15] to develop a literature review process and a review protocol that fit our purpose. Due to our study's aim to identify the key PBMs of high quality with respect to their theoretical foundations, reporting, and methodologies, we conducted a thorough paper quality assessment as part of the literature review. The steps of the process were as follows:

(1) definition of the research questions (RQs),
(2) review protocol development,
(3) literature search,
(4) study selection,
(5) study quality assessment,
(6) data extraction, and
(7) data synthesis.

To investigate whether the PBMs' differences could be explained by the various background theories and application domains used for their design, we formulated the following RQs:

> RQ1. What background theories have been applied in the existing PBMs' initial construction and subsequent development, and how has this been done?
>
> RQ2. How has the data collection domain been taken into account in the PBMs' construction and development?

RQ1 was aimed to gain insights into the background theories' usage in the modeling and obtain information on the role of information privacy concerns in the PBMs based on the applied theories. RQ2 was intended to provide information on the ways that the data collection domain had been considered in the PBMs to adapt them to different types of personal data collection contexts. Together, the answers to RQ1 and RQ2 would provide information on the reasons for the PBMs' differences in their variables' relationships.

We selected three digital databases for our search due to their high relevance to the research topic. The reference database Scopus was chosen as the main resource because of its broad coverage of high-quality publication forums. Scopus was complemented by the multidisciplinary, scholarly full-text database Academic Search Premier and by Business Source Complete, which contains articles from top-ranking journals in both business and management fields. As PBMs encompass various types of data collection contexts and data usages, our literature review's scope was limited to personal data collection for commercial transactions, service production, and other business purposes, as well as Internet usage in general. This limitation was taken into account in the design of the study selection criteria.

Prior to our actual literature search, we conducted several pilot searches with different search terms. In this way, we acquired insights into the varied factors and terminology used in PBM research and could thus better formulate the search string to cover this diversity. The final search string included the following keywords:

> privacy AND (information OR data) AND (disclos* OR collect* OR acquisition OR release OR expose OR share OR revelation) AND (concern OR risk OR cost OR benefit OR payoff OR calculus OR willing*) AND model.

"Title", "abstract," and "keywords" were used as the search fields. We conducted the main search in March 2017, which was limited to the studies published by the end of 2016. To cover more recent works published in 2017, we performed a complementary search in February 2018. Through the search, we identified a total of 1,353 studies. The SLR process is presented in the following paragraphs. Appendix A describes it in the form of a flow diagram, adapted from the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines.[16]

We conducted the study selection in three stages by applying the criteria presented in Appendix B. In the first stage, duplicates and papers with a publication forum or a content type that was irrelevant to the study were excluded based on the metadata (741 papers were included in this stage). In the second stage, the papers that did not present an empirical model describing privacy behavior were excluded based on their titles and keywords (273 papers were included). In the third stage, papers were excluded based on their abstracts if they did not present a PBM with a behavioral outcome, if the model was not based on empirical results, or if they did not study the data subjects' viewpoints (139 papers were included). One researcher handled the first stage; two researchers dealt with the second and the third stages. In the first stage, the researcher marked the papers either "accept" or "reject" based on the defined criteria. In the second and the third stages, the researchers could also choose "can't decide" as the third option. In both the second and the third stages, the decision on inclusion or exclusion was based on the researchers' combined results. In case both researchers agreed to accept or reject a paper (or only one marked the paper "can't decide"), the decision to accept or reject was made accordingly. If the researchers disagreed or both marked the paper "can't decide", it was marked as a conflict. After the second and the third stages, conflict resolution meetings were held. If the researchers could not agree after the second selection stage, the undetermined papers were moved to the third stage, along with the selected papers. If conflicts remained after the conflict resolution meeting held after the third selection stage, a third researcher would be asked to participate in the meeting for a final resolution (however, in practice, there was no need for this step).

Study quality assessment is an important step in an SLR to gather research with solid methodology and high validity. We developed a checklist based on the four aspects of quality assessments by Zhou et al.[17] and applicable questions by Kitchenham and Charters.[15] Appendix C presents our checklist questions and corresponding measurement scales. The final study set was selected based on the minimum quality threshold and the studies' citation index. The minimum quality threshold was determined to result in a manageable quantity of studies with a quality score close to full (i.e., meeting all the quality criteria). As we aimed to identify a focused set of the most often cited key papers, we also used the Field-Weighted Citation Impact (FWCI) metric for the papers' selection and excluded papers with an FWCI score lower than 1 (i.e., below the average of the subject field).

We conducted content analysis of the key papers to identify the background theories and the application domains of the models presented in these papers. As data subjects' information privacy concerns are among the key privacy behavior antecedents reflecting privacy perceptions in a holistic way (cf. information privacy concerns evaluation instruments, e.g., refs. 18–20), we also identified these concerns' conceptualizations and roles in the existing PBMs to support the analysis. We recorded the following information from each paper: the PBM's theoretical background, application domain, behavioral outcome, key privacy behavior antecedents, conceptualization of privacy concerns, and cited previous studies used as the basis for the PBM. Regarding privacy behavior antecedents, their definitions or descriptions were also recorded, as well as the definitions' sources. Based on this information, we could identify the distinct background theories used in the PBMs' construction and obtain an overall view of the models' constructs. The recorded information was used (together with the researchers' notes about the papers during the recording) for summarizing each identified PBM with respect to its theory usage, application domain, conceptualization of privacy concerns, and relationships among privacy concerns and other privacy behavior antecedents.

Results

In our literature review, we identified 11 primary studies (marked with an asterisk in the references) that each presented one PBM. Nine of these studies were conducted in the 2010s. This reflects the novelty of the research topic, as well as the currently increasing need for understanding information privacy behavior in the evolving personal data collection contexts. Next, we present the summaries of the 11 PBMs to provide an overview of their approaches to explaining privacy behavior in different types of application domains. Furthermore, we review the background theories referred to in the construction of these PBMs and explain in more detail how the theories are applied in them. We describe how the application domain is considered in the PBMs as well. Table 1 summarizes the information on our primary studies' PBMs regarding their application domains, behavioral outcomes, and key privacy behavior antecedents (the studies are arranged chronologically).

The application domains of the primary studies' PBMs covered Internet usage in general, e-commerce and other online transactions, healthcare data collection, online services, driving data collection, and telecommunication. These PBMs' behavioral outcomes were most often associated with personal data disclosure, such as behavioral intention to disclose personal information or willingness to disclose this information. It is noteworthy that behavioral outcomes other than these were also incorporated in the models. These were associated with providing access to personal data or authorizing their secondary usage, information misrepresentation, and loyalty toward a technology. The majority of the PBMs included privacy concerns, risk-related variables, and trust-related variables as privacy behavior antecedents. For their part, benefits were included in six PBMs.

Summaries of the identified PBMs

In the summaries of the primary studies' PBMs, we briefly describe how each PBM explains privacy behavior in its particular study domain. We describe the PBMs' application domains, theory usage, model constructs, relationships among the constructs, the roles of privacy concerns, and the ways in which privacy concerns are defined.

P01

The presented model explains the relationship between Internet users' information privacy concerns and behavioral intention toward releasing their personal information to an online marketer.[4] The model is grounded on the trust-risk model and the theory of reasoned action (TRA)—trusting and risk beliefs associated with personal information release are incorporated into a TRA-type model with the behavioral intention component of releasing information. The authors consider privacy concerns an individual characteristic that affects these beliefs and—mediated by them—behavioral intentions. They note that both the trust-risk literature and the TRA support this kind of relationship. The notion of privacy concerns is characterized in terms of three factors based on the social contract (SC) theory: information collection, control over the information, and awareness of privacy practices. Information privacy concerns are defined as an individual's subjective views on fairness in the context of information privacy, following Campbell's study.[27] It is pointed out that privacy concerns are influenced by external conditions, such as industry sector, culture, and legislation. On the other hand, based on Donaldson and Dunfee's study,[28] the authors state that individuals' perceptions of these conditions vary with their personal characteristics and past experiences.

P02

The presented model focuses on the data subjects' beliefs about privacy and their corresponding behavioral intentions in the Internet transaction domain.[5] The model derives from the TRA and its later version, the theory of planned behavior (TPB). It incorporates two primary components of the TRA and the TPB models—beliefs and behavioral intention—to investigate the beliefs that influence the behavioral intention to disclose the personal information needed for Internet transactions. The authors assume that these beliefs can be contradictory by nature and together comprise a set of elements in a data subject's privacy calculus-type decision

Table 1. Application domains, behavioral outcomes, and key privacy behavior antecedents of primary studies' privacy behavior models (PBMs).

				Key antecedents included in the study			
Primary study	Index	Application domain	Behavioral outcome	Privacy concerns	Risk-related variables	Trust-related variables	Benefits
4	P01	Internet (online marketing)	Behavioral intention to release personal information	x	x	x	
5	P02	E-commerce	Willingness to provide personal information	x	x	x	
10	P03	Healthcare data collection	Willingness to provide access to personal health information	x	x	x	
21	P04	Online services	Willingness to disclose personal information	x	x	x	x
22	P05	Commercial websites	Behavioral intention to disclose personal information	x		x	x
23	P06	E-commerce	Behavioral intention to disclose personal information	x		x	
9	P07	Smartphone application with driving style feedback	Behavioral intention to disclose personal information	x	x	x	x
24	P08	E-commerce	Behavioral intention to provide personal information, information misrepresentation	x		x	x
25	P09	Telecommunication companies' customer data collection	Behavioral intention to authorize the secondary use of personal information	x	x	x	x
12	P10	Online business transactions (finance and e-commerce)	Behavioral intention to disclose personal information	x	x	x	
26	P11	Mobile hotel booking (MHB)	Loyalty intentions toward MHB technology	x	x	x	x

process (i.e., their cost-benefit evaluation of their personal information disclosure) that leads to one's intention to disclose personal information to complete a transaction. Overall, the model is grounded on the expectancy theory; data subjects are assumed to behave in ways that maximize positive outcomes and minimize negative outcomes of their behavior. Following the privacy calculus theory, risk beliefs and confidence and enticement beliefs are incorporated into the model. A direct relation is suggested between two risk beliefs: perceived Internet privacy risks and privacy concerns. In this model, privacy concerns are considered an *internalization of the possibility of privacy loss* associated with websites in general. Such concerns represent an assessment about what happens to the personal information that the user discloses on the Internet. When defining their model construct for privacy concerns, the authors also refer to the *possibility of other parties' opportunistic behavior* related to the submitted personal information. They particularize the concept of privacy concerns by stating that these comprise the data subject's beliefs about who has access to the disclosed information and how it is used.

P03

The presented model extends the privacy calculus and its consequentialist approach to privacy in the healthcare data collection domain.[10] Through this extension, it aims to take into account privacy risks associated with sensitive health information, as well as emotions linked to the data subject's health condition. The authors augment the privacy calculus approach with the communication privacy management (CPM) theory to deal with the data collection context. In this way, they incorporate situational risk factors, specific to the context, into the model. The authors also apply the risk-as-feelings perspective to deal with emotions linked to a person's health condition. Both the situational risk factors and the emotions are considered key antecedents of the willingness to disclose information. Situational risk factors affect such willingness by moderating privacy concerns (no definition of privacy concerns could be found in the paper), whereas emotions directly affect this behavioral outcome. Compared with cognitive evaluations, emotions can result in insensitivity to probability variations in risk perceptions. For example, an individual can focus more on the desire to improve one's health and feel better when disclosing information despite the lower probability for health improvement compared with the potential privacy risk.

P04

The presented model explains the role of information sensitivity in personal information disclosure for online services, as well as its effects on online privacy concerns, perceived customization benefits, and perceived information control.[21] The authors base their model on the prospect theory. Through this theory and its value function concept, they study the differential effects of disclosure antecedents on willingness to disclose as a function of information sensitivity. Therefore, in their model, online privacy concerns' negative influence on the willingness to disclose information online is stronger for highly sensitive information. Furthermore, the positive effect of perceived customization benefits on the willingness to disclose information is stronger for less sensitive versus highly sensitive information in the presence of greater online privacy concerns (not in the presence of less privacy concerns). The model includes the construct of perceived information control that is often considered closely related to privacy concerns. Contrary to the common approach, control is defined as a distinct construct that affects the willingness to disclose information. The authors refer to online privacy concerns as consumer concerns about the use of their disclosed information for marketing purposes, beyond its intended purposes of use. These privacy concerns are considered personal dispositions based on the studies of Malhotra et al.[4] and Son and Kim.[29]

P05

The presented model explains the effects of positive and negative affects on users' trust and privacy beliefs (also referred to as privacy concerns), as well as their intentions to disclose personal information to commercial websites.[22] The author bases the model on the congruity theory due to this theory's capability of dealing with communication and persuasion contexts and changes in individuals' cognitions. In this model, users who visit an unfamiliar e-commerce website maintain trust and privacy beliefs that are consistent with their general Internet security beliefs. However, website cues, such as their friendliness, helpfulness, and enjoyment experienced by users, may change users' attitudes toward a website by altering their affective state. In this way, the cues have an effect on the intention to disclose information, both directly and mediated by the website trust and privacy beliefs. Through these relationships between variables, the author aims to contribute to the privacy paradox theory development (i.e., data subjects disclosing personal data despite their high level of privacy concerns). In the model, website privacy beliefs are considered antecedents of website trust. Such beliefs are characterized as targeted perceptions about specific elements of the Internet. Referring to the study of Hoffman et al.,[30] the author states that these beliefs exist when personally identifiable information related to an individual is collected and stored in digital format, and they stem from individuals' lack of awareness about the usage and the distribution of this information.

P06

The multilevel model of information privacy beliefs presented in the paper explains the effects of distinct privacy beliefs on the intentions to disclose information and make a transaction on a particular e-commerce website.[23] The author draws from the earlier literature that suggests privacy as a context-dependent concept. In this view, data subjects' experiences with the general online environment may not extend to all websites but differ as a result of specific websites shaping their perceptions. The author bases his model on the attitude theory and the TRA to deal with the data subjects' multiple attitudes and beliefs across contexts, as well as possible changes in these attitudes and beliefs through persuasion and social influence associated with a specific context. The model distinguishes among three levels of privacy beliefs: disposition to privacy, online privacy concern, and website privacy concern.

Disposition to privacy affects both online and website privacy concerns, and online privacy concern influences website privacy concern. Together, these privacy beliefs affect the data subjects' behavioral intentions. The effect of disposition to privacy is attenuated and mediated by more contextualized privacy beliefs (i.e., online and website privacy concerns). Moreover, in this model, the effect of website privacy concern on behavioral intention is partially mediated by trusting beliefs. The author characterizes disposition to privacy as a personality trait based on the work of Xu et al.[31] Following the studies of Dinev and Hart[5] and Malhotra et al.,[4] he defines online privacy concerns as beliefs that reflect a person's overall perception of privacy risks. For their part, website privacy concerns are defined as situation-specific privacy beliefs based on previous research.[1,32,33] These concerns deal with uncertainties regarding private information handling by a particular website.

P07

This paper explains the data subjects' intention to disclose their driving behavior data for a smartphone application with customized feedback on their driving styles.[9] The authors develop an extension of the privacy calculus-type model through the affect heuristic approach from consumer behavior research to study the situation-specific tradeoff between privacy risks and data disclosure benefits. In this way, they take into account not only data subjects' dispositional factors (general privacy concerns and general institutional trust are considered such in this study) but also their limited cognitive resources, heuristic thinking, and affective reactions in situation-specific behavior. The model is based on the assumption that the data subjects' situation-specific risk perceptions, associated with their information disclosure in a particular data collection context, may override their dispositional tendency to worry about information privacy. For this reason, the data subjects' assessment of situation-specific risks may persuade them to disclose their information despite general privacy concerns. In this model, general privacy concerns are considered an antecedent of a situation-specific risk assessment. The authors use Li et al.'s[34] definition of privacy concerns as an individual's tendency to worry about information privacy.

P08

The presented model explains the effects of monetary rewards on information privacy concerns, personal information disclosure, and information misrepresentation on e-commerce websites requesting information with different sensitivity levels.[24] The model rests on the privacy calculus theory, suggesting that data subjects evaluate the cost of personal information disclosure against the received compensation. An interaction effect is expected between monetary rewards and information sensitivity on information privacy concerns and behavioral intentions. Contrary to the common idea of the cost-benefit evaluation, it is assumed that monetary rewards may also stimulate the data subjects' risk perceptions and elevate their concerns about disclosing sensitive information. The authors associate privacy concerns with the data subjects' perceived unfair loss of control over their privacy. Following the studies of Malhotra et al.[4] and Donaldson and Dunfee[28], they highlight the subjective nature of privacy concerns that can be considered "an individual's subjective views to fairness within the context of information privacy" and stemming from an individual's own perception and values. The authors also refer to the study of Malhotra et al. regarding the effect of environmental factors (e.g., industry, culture, and legislation) on privacy concerns. Furthermore, they draw parallels between privacy concerns and information sensitivity, following Weible's definition of information sensitivity as "the level of privacy concern an individual feels for a type of data in a specific situation."[35] The authors suggest that information sensitivity and information privacy concerns are conceptually similar because one can be defined by the other, and both account for subjective risk perceptions.

P09

The presented model explains the role of information practices on consumers' intention to authorize the secondary use of their personal data by telecommunication companies.[25] The authors base their model on the TRA, the CPM, and the expectancy theory. They incorporate trusting beliefs and, motivated by the CPM, perceived information risks and benefits from personal data disclosure into a TRA-type model. Authorization of the secondary use of personal data is incorporated into this model as a behavioral intention component. The roles of perceived benefits and information risks in the model are grounded on the expectancy theory. In the model, information privacy concerns' effect on behavioral intention is fully mediated by behavioral beliefs (i.e., trusting beliefs and risk perceptions). Similarly, information privacy concerns' effect on behavioral intention is mediated by the data subjects' perceptions about information practices. The authors draw parallels between information privacy concerns and personality traits (based on ref. 4), and define information privacy concerns as individuals' worries about the possible loss of information privacy (based on ref. 31). They also state that these concerns are the data subjects' subjective evaluations, possibly varying with individuals' personal characteristics and past experiences (based on ref. 13).

P10

The presented model explains the roles of the data collection context's sensitivity and the customers' personalities in their disclosure behavior in online business transactions.[12] The authors base their model on the TRA. Similar to the technology acceptance model (TAM), they simplify the TRA by excluding its subjective norm component. On these grounds, they incorporate privacy concerns into the model as a dispositional belief that affects behavioral intention, both directly and as mediated by trust (the latter relation is established following the study of Malhotra et al.[4]). The authors synthesize the TRA with the prospect theory to contextualize their model so that it takes into account data subjects' personality traits, previous experiences with privacy invasion, and the context's sensitivity. Based on this, in the model, privacy concerns are influenced by the data subjects' personality traits and previous experiences with online privacy invasion. For its part, context sensitivity is treated as a moderator of the variables' relationships. The authors consider privacy concerns as constituting the salient dispositional belief, drawing from the TRA.

P11

The presented model explains mobile hotel booking (MHB) users' loyalty intentions toward MHB technology with personalized services.[26] The authors base their model on the privacy calculus theory. However, they state that the privacy calculus alone may be insufficient for this type of study—privacy concerns and personalization benefits are not necessarily the only determinants of the data subjects' behavior. The authors integrate the privacy calculus and the trust-risk model to examine a wider set of the determinants of loyalty intentions, including data subjects' trust in the service provider and risk perceptions about the MHB technology (these are not privacy risks as such). In the model, privacy concerns affect loyalty, both directly and mediated by trust, and are consecutively mediated by trust and perceived risk. In turn, privacy concerns are influenced by personalization of services. The authors take into account the unique features of mobile commerce data collection, including portable personal items and the possibilities that service providers will trace users' activities. They refer to the four dimensions of privacy concerns identified by Smith et al.,[18] but focus on the dimensions of location information usage, data collection, and secondary use of information to deal better with the mobile commerce domain.

Background theories and their application in PBMs

We identified 13 background theories that were referred to in the construction of the presented PBMs. In each primary study, at least one theory was used for the model construction. Typically, two or more background theories served as starting points for the presented PBMs. The theories referred to in the primary studies are presented in Table 2 (the studies are arranged chronologically).

Table 2 shows that six of the theories (in bold) were referred to by two or more studies. The other seven theories were each referred to only in one study. Of the more frequently referred theories, the *TRA* was found in five studies (P01, P02, P06, P09, and P10). Four studies (P02, P03, P07, and P08) based their PBMs on *privacy calculus thinking*. Another behavioral economics theory, the *prospect theory*, was used as the basis for the PBM in P04 and synthesized with the TRA in P10. Two studies (P02 and P09) referred to the *expectancy theory*, two (P03 and P09) to the *CPM theory*, and two (P01 and P11) to the *trust-risk model* as well.

To assess the representativeness of our identified theory set, we compared it with the results of an earlier theory review by Li.[13] The 13 background theories that we identified appeared to correspond fairly well to the results (15 theories) of this earlier review. Li's theory set did not include seven of the theories that we identified, as follows: prospect theory, congruity theory, trust-risk model, attitude theory, risk-as-feelings, TAM, and theory of affect heuristic. Five of these theories were each referred to only once in our primary studies, whereas the prospect theory and the trust-risk model were referred to twice. Our theory collection did not contain nine of the theories mentioned in the earlier review, as follows: agency, utility maximization, expectancy value, procedural fairness, social presence, social response, protection motivation, social cognitive, and personality theories. However, in Li's study, none of these theories was among the most referred ones. Based on the comparison, the most often referred theories that we identified appeared to be the same as those in Li's review. The correspondence of our theory set to that of the earlier privacy behavior model review suggests that our set represents relatively well the theories referred to in PBM research.

In Table 3, we summarize the general ideas of the six most often referred background theories with respect to their original application areas and core principles. We then briefly review together the theories that were referred to only once among the studies.

Seven theories that were each referred to only in a single study were used for model construction, either to support or complement other background theories' presuppositions, or to contextualize the PBM in a specific data collection environment. Of these theories, the *TPB*, the *congruity theory*, and the *TAM* deal with attitudes as the antecedents of behavioral intention, similar to the TRA. They also draw from the same theoretical tradition as that of the TRA. For its part, the *attitude theory* of social psychology explains the change in

Table 2. Background theories referred to in the primary studies.

		Background theory												
Primary study	Index	**TRA**	**Expectancy theory**	**CPM**	**Privacy calculus**	**Prospect theory**	**Trust-risk model**	TPB	Congruity theory	SC theory	Attitude theory	Risk-as-feelings	TAM	Theory of affect heuristic
4	P01	x					x			x				
5	P02	x	x		x			x						
10	P03			x	x							x		
21	P04					x								
22	P05								x					
23*	P06	x									x			
9	P07				x									x
24	P08				x									
25	P09	x	x	x										
12	P10	x				x							x	
26	P11				x		x							
Total number of references		**5**	**2**	**2**	**5**	**2**	**2**	1	1	1	1	1	1	1

*TRA is not directly mentioned, but based on the references and the presented model, it is used as part of the theoretical background.

Table 3. Summaries of the background theories referred to in the primary studies.

Theory	Summary	References
Theory of reasoned action (TRA)	The TRA is a general theory of behavior in social psychology, grounded on an individual's basic motivation to demonstrate a behavior. The individual's behavioral intention determines his or her actual behavior. Behavioral intention is determined by the individual's attitudes toward a behavior, with respect to its outcome, his or her subjective norms (i.e., the perceived social pressure to show or not to show the behavior), and beliefs that affect these variables.	36,37
Expectancy theory	The expectancy theory is a psychological theory of motivation suggesting that individuals seek to maximize positive outcomes and minimize negative outcomes of their behavior. An individual decides whether to perform an action through a cognitive process with three elements: the individual's expectation that a specific effort will lead to the intended performance, the instrumentality of this performance in achieving the desired result, and the desirability of the result in question for the individual.	38
Communication privacy management (CPM) theory	The CPM theory is a social psychology theory of individuals' decision-making about either revealing or concealing private information specifically in interpersonal situations. Individuals control their privacy based on the expected costs and benefits of information disclosure. This privacy management is illustrated through a metaphor of privacy boundaries between individuals and their communication partners. Individuals govern their information disclosure in a rule-based manner through these boundaries.	39
Privacy calculus	The privacy calculus is a behavioral economics theory derived from the "calculus of behavior." In privacy calculus, an individual's tradeoff between positive and negative consequences of personal information disclosure determines his or her intention to disclose this information (i.e., behavioral intention depends on both costs and benefits of information sharing). The traditional privacy calculus involves the presupposition of rational decision-making, hence ignoring individuals' incomplete information, bounded rationality, and bounded cognitive ability to process the information.	3,40–42
Prospect theory	The prospect theory is a behavioral economics theory that describes individuals' decision-making with respect to probabilistic alternatives involving risks. These decisions are based on the value of potential gains and losses associated with the expected outcomes (i.e., individuals' expected utility and disutility). The theory suggests individuals' loss aversion through its value function (i.e., the negative feeling of loss is greater than the positive feeling of equivalent winning). In contrast to the traditional privacy calculus, it emphasizes the role of heuristics in decision-making.	43,44
Trust-risk model	The trust-risk model is based on the research conducted in the field of organizational relationships. It explains the roles of interpersonal trust and experienced risks in individuals' behavior in uncertain situations. The effect of personality traits on trust is also often incorporated into the trust-risk model.	45,46

individuals' attitudes by persuasion and social influence. The *theory of affect heuristic* and the *risk-as-feelings perspective* of psychology explain the role of emotional responses in decision-making. The *SC theory* from political philosophy is applied to business contexts when studying individuals' perceptions of fairness and justice. In this way, it can be used to explain the consumer–firm relationship.

In conclusion, it can be considered that the identified theories fall under two main types with somewhat distinct grounds. The majority of the primary studies' PBMs are derived from theories that present individuals' expectations of outcomes of their behavior (e.g., data disclosure in PBMs) as their behavior antecedents. These theories explain how individuals' evaluations of the expected outcomes affect their decision-making and through this, their behavior (TRA, expectancy theory, and trust-risk model). Individuals' evaluations of the expected outcomes and the corresponding decision-making can be compared with their risk assessment (i.e., evaluation of uncertainty and severity of outcomes of activities) (cf. ref. 47). In contrast to these outcome evaluation-based theories, the privacy calculus, the prospect theory, and the CPM theory are based on individuals' quantified cost-benefit evaluation of the value resulting from their behavior. This evaluation is guided by the individuals' perceived net value of the outcomes and their preferences, instead of the expected outcomes as such, and can be perceived as more subjective than the outcome-based evaluation in privacy behavior. Of the theories that were referred to only once, some follow the same classification—the TPB, the congruity theory, and the TAM deal with the expected outcomes of certain behaviors.

Overall, many of the identified theories are relatively generic by nature. The TRA, the TPB, and the congruity theory are developed to explain the relationships among attitude, behavioral intention, and performance in general, instead of privacy behavior in personal data collection domains. Similarly, the expectancy theory, the prospect theory, and the trust-risk model deal with behavioral decision-making in general. Contrary to these, the privacy calculus and the CPM specialize in information privacy and personal information disclosure contexts.

Among the primary studies, typically two or more background theories were jointly used for the model construction, and only three studies were based on a single theory (P04, P05, and P08). It was often not possible to explicitly identify any principal background theory when multiple theories were referred to. Instead, distinct theories were typically conjoined in a mutually enhancing way to design and develop a solid PBM. Moreover, the two types of background theories described in a previous paragraph (outcome evaluation-based theories and cost-benefit evaluation-based theories) were used in parallel in a few studies (P02, P09, and P10). For example, the PBM presented in P02 was based on the TRA and the TPB, together with the privacy calculus and the expectancy theory. It is noteworthy that the theories referred to in the primary studies were really used for designing their models, contrary to just discussing these theories generally (studies of this type were identified in the quality assessment phase of our literature review).

In all studies drawing from the **TRA**, this theory was referred to together with other background theories (P01, P02, P06, P09, and P10). All these studies incorporated the behavioral intention component of the TRA in their models.

Otherwise, the primary studies showed considerable diversity in their application of the TRA components, and typically, not all of the TRA components (specifically, subjective norms and normative beliefs) were incorporated into the models. Such components were excluded from the models by referring to the earlier literature that focuses on some core components of the theory (P01 and P09) or by explaining the included components' relevance to the studies. Privacy behavior antecedents (e.g., privacy concerns, perceived information disclosure risks, and trust) were often incorporated into the model as beliefs affecting the data subjects' behavioral intentions. The attitude component was included only in one model (P10). In the model presented in P06, the TRA played a minor role. However, when distinguishing general beliefs from more specific beliefs, based on the TRA, this study stated that an attitude's power to predict a behavior depends on how closely the attitude relates to the behavior.

Similar to the TRA, the **expectancy theory** was used in parallel with other theories for the model construction in two studies. This theory was employed in P02 and P09 to deal with the expected negative and positive outcomes of a behavior. In both studies, information disclosure risk perceptions were considered representative of the negative outcomes. Correspondingly, consistent with the expectancy theory, both studies hypothesized the association between high-risk perceptions and the behavioral intention to withhold information. In P09, perceived benefits (of both monetary and non-monetary types) of personal data usage authorization were considered positive outcomes of this behavior. P02 also drew from the privacy calculus, and in its model, benefits were represented by the data subjects' confidence and enticement beliefs.

Similar to the TRA and the expectancy theory, the **CPM** was referred to in combination with other theories in PBM construction (P03 and P09). In P03, the CPM was utilized together with the health informatics literature to study individuals' personal health information disclosure. Based on the CPM, it was suggested that individuals erect boundaries differently around various types of personal health information. In P09, the CPM was referred to with respect to its established roles of perceived benefits and information risks in individuals' decision-making about either revealing or concealing personal information. However, the study did not provide a more detailed description of how this theory was applied in the model construction to explain information practices' role in the data subjects' intention to authorize secondary use of their personal data.

The **privacy calculus** was referred to in five studies (P08, P02, P03, P07, and P11) and utilized in different ways for their model construction. The model presented in P08 was based solely on the privacy calculus theory. This study on monetary rewards' effect on privacy behavior rested on the idea that the data subjects evaluate the cost of personal information disclosure against the received compensation in their privacy calculation. In P02, *risk beliefs* and *confidence and enticement beliefs* were incorporated into the model, adopted from the belief component of the TRA and the TPB. These beliefs were considered comparable with privacy calculus variables, and the importance of personal beliefs was highlighted as part of the privacy calculus because individuals often cannot predict the future outcomes that they aim to manage. The model presented in P03 was based on the privacy calculus but contextualized in the healthcare domain. In P07, the traditional privacy calculus was complemented with the *theory of affect heuristic* to deal with the data subjects' limited cognitive resources, heuristic thinking, and affective reactions in their decision-making. This approach was considered a way to cope with large deviations in individuals' privacy concerns and behavioral intentions found in earlier research. In P11, the privacy calculus was integrated with the trust-risk framework to examine users' loyalty to service providers in personalized MHB services. In this way, the effects of the MHB users' trust in the service providers, as well as their risk perceptions of the MHB technology, were taken into account, in addition to their privacy concerns and perceived value of personalized services. The presented model was based on the assumption that the value of personalized services does not outweigh the loss of privacy associated with personal data disclosure. For this reason, service personalization was expected to positively affect privacy concerns.

The **prospect theory** was referred to in P04 and P10. In P04, the theory was used as the sole basis for the PBM to examine the disclosure antecedents' differential effects with respect to the requested information's sensitivity. In P10, the TRA was synthesized with the prospect theory to deal with data subjects' personalities and previous experiences with privacy invasions (based on the idea that individuals differ in their utility functions). Similar to P04, P10 examined the sensitivity of the data collection context through the prospect theory.

The **trust-risk model** was referred to in P01 and P11. In both studies, this theory was used for model construction, together with other theories. In P01, the trust-risk model was utilized in combination with the TRA and the SC theory. Drawing on the TRA, trust and risk were incorporated into the model as trust and risk beliefs. Privacy concerns were considered a data subject's individual characteristic that affects these beliefs. In P11, the trust-risk model was integrated with the privacy calculus. In this way, the authors complemented the model with trust and risk variables, in addition to privacy calculus variables (i.e., information privacy concerns and benefits from the data disclosure), to explain users' loyalty to MHB service providers.

The seven identified theories that were referred to only once were typically utilized in model construction in parallel with other theories (except for the congruity theory referred to in P05). These theories contributed to a model's development by either complementing its other theoretical backgrounds with relatively generic knowledge on individuals' behavior (TPB, TAM, attitude theory, and SC theory) or providing the model with certain aspects specific to the studied data collection context (theory of affect heuristic and risk-as-feelings).

As stated earlier in this section, many of the primary studies' background theories (e.g., TRA, expectancy theory, and prospect theory) are relatively generic by nature. They neither focus on information privacy nor really include privacy behavior variables as their model constructs. Contrary to these theories, the privacy calculus and the CPM specialize in privacy and personal information disclosure contexts. However, it seems that despite their focus on information privacy, they provide relatively loose starting points for the PBMs' development because they do not explicitly define the roles and the relationships among different privacy behavior variables.

In the PBMs constructed on these grounds, privacy behavior's key concepts can be defined in various ways, even in models with the same theoretical background, as the theories do not necessarily provide any specified definitions of these concepts. The majority of the privacy concerns' definitions that we identified were based on a diverse body of previous theoretical or empirical literature. These definitions varied in their views on privacy (i.e., the data subject's view and the data collector's view), levels of subjectivity, and application specificity. Similar to privacy concerns' definitions, their antecedents (e.g., dispositional tendency) can be defined in various ways based on the findings of previous studies and additional theories other than the background theories.

Overall, PBM construction and development seem to have been undertaken in an eclectic way so far, building on different theories and earlier empirical research. Using several background theories makes it possible to define model constructs and their relationships in diverse ways, which possibly leads to model inconsistencies.

Application domains of the PBMs

Based on our primary study set, it seems that despite the relatively generic nature of the theories referred to in PBM development, they can be used as the basis of privacy behavior modeling in different types of personal data collection domains. Theories have been applied as such by carefully selecting one that is appropriate to the domain in question. Theories have also been adapted to fit distinct domains or applications by defining new model constructs and their corresponding measurement items. The personal data collection domain was taken into account in either one or both of these ways in nine studies.

In four studies, the personal data collection domain was considered primarily when selecting the theoretical background of the presented model (P02, P03, P06, and P11). In P02, the privacy calculus was employed as a starting point for the study on online transactions, referring to Culnan and Armstrong [3], who stated its suitability for purchase contexts in general. The privacy calculus was then adapted to the studied domain by selecting and defining the model constructs consistently with it. For their part, the authors of P03 aimed to extend the traditional privacy calculus to the personal health information collection domain. They contextualized their model through situational risk factors (type of requested information, intended purpose of information use, and requesting stakeholder) and by utilizing the risk-as-feeling hypothesis, the emotion linked to the data subject's health condition. In P06, the TRA was combined with the *attitude theory* to model privacy concerns across contexts. Through this, three levels of privacy beliefs were specified: the data subject's disposition to privacy, online privacy concerns, and website privacy concerns. Of these beliefs, website privacy concerns represented situation-specific privacy concerns associated with a particular website. The authors of P11 doubted the sufficiency of the traditional privacy calculus in the domain of personalized MHB services. The authors emphasized the role of users' trust and risk perceptions in their loyalty to personalized MHBs and to investigate this, integrated the privacy calculus with the trust-risk framework.

In the models presented in P01 and P04, the domain was taken into account through the model construction and the research setting. The authors of P01 stated that only a few studies dealt with the effect of information sensitivity on the data subjects' information disclosure. Motivated by this gap, they incorporated the type of the collected information as a contextual variable into their model (their questionnaire included two scenarios of personal data collection with different levels of information sensitivity). Likewise, in P04, the effect of information sensitivity on privacy behavior was investigated in the online service domain, together with privacy concerns, perceived control over personal information usage, and perceived customization benefits. Information sensitivity was incorporated into the model as a moderator of the relationships between privacy behavior antecedents and the willingness to disclose information. For their part, the authors of P08 considered information sensitivity a variable that affects privacy behavior separately from the data collection context. However, they based their model on the assumption that monetary rewards for personal data disclosure either mitigate or intensify privacy concerns, depending on the data collection context, and this effect moderates the influence of information sensitivity.

In P10 and P07, the data collection domain was considered by both selecting their models' theoretical backgrounds and inserting new constructs in them. In P10, the TRA was used as the starting point for the study by contextualizing and synthesizing it with the prospect theory to develop the TRA–Privacy theory. Through this TRA–Privacy theory, the effect of the data collection context's sensitivity on privacy behavior was studied in online business transactions. To study data collection domains of different types, three domains with different information sensitivity levels (finance, e-commerce, and health) were included in the laboratory experiment as the moderators of the TRA paths. Similar to P02 and P03, the authors of P07 complemented the traditional privacy calculus in the mobile data collection domain. To deal with the data subjects' limited cognitive resources and affective reactions in their decision-making, they combined the privacy calculus with the theory of affect heuristic. In this way, they introduced a situation-specific privacy calculus for a smartphone application that collected driving behavior data. The authors incorporated a situation-specific assessment of risks and benefits into their model, as part of this extension of the privacy calculus, to deal with the domain with highly sensitive consumer information collection and risks of data misuse. According to the model, this assessment is affected by the collected data's sensitivity and the data subject's affective state, depending on the user interface's persuasive characteristics, and it may override dispositional privacy concerns.

In their studies, the authors of P05 and P09 focused only on the privacy behavior variables that were not dependent on the personal data collection domain. For this reason, they did not consider the domain aspect in selecting their models' theoretical backgrounds or constructs.

Our findings indicate the possibly increasing tendency in PBM development to adapt models to different personal data collection domains. This adaptation can be achieved by integrating more specific theories from psychological research or behavioral economics, for example. Moreover, the collected data type (particularly information sensitivity) has often been incorporated into a model as a distinct construct to study the effect of the data collection domain on privacy behavior.

Discussion

Our study's objective was to gain an understanding of the differences among the PBMs by investigating whether these differences could be explained by the PBMs' background theories and their application domains. We reviewed and analyzed the background theories and their usage in the existing PBMs' construction and development, as well as investigated how application domains were considered in them. In our analysis, we focused on information privacy concerns and their conceptualizations and roles in the PBMs. Through this research, our study provides insights into PBM construction and development in different types of data collection domains, as well as the possible need for further development of the currently used PBMs' background theories.

Implications for theory

Among our primary studies, we identified six relatively generic theories that have been continuously referred to in PBM development. These theories can be classified into two categories based on their criteria for individuals' decision-making: outcome evaluation-based and cost-benefit evaluation-based theories. The identified theories were typically used for PBM construction in combination with other theories. However, any principal background theories were not identified as such; rather, distinct theories were conjoined in an eclectic way for a specific purpose and *ad hoc*-type modeling.

Our analysis indicates that a PBM typically builds on one to four background theories with various assumptions and definitions of privacy behavior's key concepts (e.g., information privacy concerns). These theories have commonly been applied in varied ways to construct PBMs. On the other hand, since many background theories are usually relatively generic and not originally developed for information privacy behavior, they do not necessarily take into account all aspects of privacy behavior as such. As a result, the way that information privacy concerns are presented as part of the PBMs is often not explicitly based on any solid theoretical background, and their roles may vary substantially among distinct models (e.g., privacy concerns may be considered the antecedents of risk perceptions or vice versa). Contrary to our assumption, it seems that the differences among the PBMs cannot be directly explained by their theoretical backgrounds. The classification of the theories into outcome evaluation-based and cost-benefit evaluation-based types does not explain the models' differences, either. Instead, the eclectic use of theories and *ad hoc*-type modeling possibly account for the existing PBMs' differences and conflicts in their constructs and the constructs' relationships. Rather than background theories as such, similarities among the PBMs could possibly be explained by their common references to earlier empirical research.

When designing a new PBM based on generic theories, these theories can be adapted to specific purposes, as well as the personal data collection domain in question, to validly describe the data subjects' behavior. This adaptation has often been achieved by integrating generic theories with other theories. In this kind of integration, some original components of a generic theory (e.g., subjective norms when applying the TRA) are typically excluded from the model (or correspondingly, new components are added). If a generic background theory is applied to a PBM in this way, we recommend carefully considering these components' relevance to the study and ensuring that the resulting model still aligns with the theory. Theories that specialize in information privacy may also have to be adapted to a particular study and its domain because these theories do not always explicitly define the roles and the relationships among different privacy behavior variables. Rather, such theories implicitly define the types of these variables (e.g., data subjects' perceived costs and benefits of personal data disclosure).

The personal data collection domain is often considered in PBM construction. Our analysis shows that the existing models have been adapted to particular domains, either through background theory selection or incorporation of contextual factors as model constructs. It seems that the integration of commonly referred background theories and appropriate theories from psychological research or behavioral economics, for example, could be utilized to contextualize a PBM in the domains where the data subjects' privacy behavior is very specific to the domains in question and possibly affected by emotions or irrational thinking, such as personal healthcare involving highly sensitive information or mobility data-based services with continuous monitoring. In the PBMs, the data collection domain is often represented by the collected data type. Specifically, information sensitivity has been incorporated into the models as a construct that affects other privacy behavior variables. However, other domain-specific factors should possibly be considered in modeling. For example, the purposes of the collected data usage and the data collectors' characteristics are aspects associated with personal data disclosure that the data subjects probably take into account in their decision-making about the disclosure. Overall, it seems that the PBMs' differences cannot be explained by the application domain. For example, differences in the PBM variables and the variables' relationships are observed in the PBMs that are constructed for the e-commerce domain, as well as when comparing these with the models with different domains. The differences are rather derived from the *ad hoc*-type model contextualization, similar to the use of background theories in PBM construction. Although PBMs have been constructed for different domains, their contextualization has not been based on a detailed analysis of the domain characteristics, from the information privacy perspective. In this respect, the research is either not yet mature or has not been properly structured.

Currently, the knowledge on privacy behavior has accumulated through theories that typically explain particular areas of the phenomenon. In fact, it is known that not all of these theories are consistent with the data subjects' actual behavior in the present complex data collection environments, such as the privacy calculus with its assumptions on rational decision-making (cf. refs. 41 and 42). Overall, it seems that PBM research is still evolving as it is drawing from diverse theoretical backgrounds, and the PBMs are typically developed by combining these theories and the results of earlier empirical research. This current state can be regarded as equivalent to Kuhn's[48] pre-paradigmatic stage of science with multiple assumptions, hypotheses, and concepts. According to Kuhn, the pre-paradigmatic stage is followed by normal science that has a certain settled paradigm for conducting research and further development of the theoretical background.

Our findings contribute to the research on privacy behavior modeling and theory development. If privacy behavior is studied inductively and in an eclectic way (as conducted so far), subsequent research will possibly result in a continuum of disjointed and conflicting PBMs. Therefore, attention should now be paid to the appropriate and well-considered application of background theories and the development of guidelines for this. We identified a wide set of PBMs' background theories and provided insights into their usage in privacy behavior modeling. These theories and their application in the information privacy domain should be analyzed in detail to compare and explicitly explain their differences with respect to their views about individuals' behavior, as well as the roles, distinctiveness, and conceptualization of privacy behavior antecedents (e.g., whether the antecedents deal with expectancies, attitudes, perceptions, cognitions, emotions, etc.). This requires a conceptual theoretical research approach and still demands substantial effort. Another future research challenge is how to analyze PBMs' differences systematically and exactly, with their complexity, divergent construct structures, and eclectic nature. It should be considered whether there is a need for structuring privacy behavior research, as well as a demand and possibilities for the development of more comprehensive and elaborate theories that are consistent with one another. Similarly, it is highly relevant for privacy research to analyze and develop model contextualization for different personal data collection domains and make privacy behavior conceptualization more consistent. Ideally, a comprehensive research framework could be constructed to be corroborated, validated, and finally applied to PBM research in different types of application domains. This type of research requires an extensive analysis of the existing PBMs and can be based on our findings about the key papers' analysis.

Implications for practice

The PBMs can be utilized in the design and the development of personal data-based systems and services to meet users' privacy requirements in a better way. However, our analysis shows that some issues should be borne in mind when applying the results of the existing PBM studies. These issues are associated with the interpretation of the results and the privacy behavior measurements.

Due to the existing PBMs' diversity, it may not be easy to identify appropriate references to gain an understanding of privacy behavior for the development of personal data-based services. Our study provides knowledge on the PBMs' theoretical and domain-specific grounds, as well as their application in the model design. This information helps in identifying studies that match personal data collection and data subjects' behavior in a specific situation. As the conceptualizations of privacy behavior antecedents may vary from one study to another, it is also advisable to carefully examine the definitions used in a particular study when applying its results. If the antecedents' definitions or descriptions are not presented, the results cannot be explicitly interpreted.

The PBM measurement items can be used for practical purposes in designing and developing personal data-based services, such as when conducting user surveys (e.g., on data subjects' information privacy concerns) or consulting data subjects about their views regarding privacy to prepare sector-specific codes of conduct under the European Union General Data Protection Regulation. The measurement items often need to be tailored to a specific situation to gain a relevant understanding of privacy behavior. The knowledge about background theories and conceptualizations of privacy behavior may be useful in this task because it serves as a guide in considering relationships among different behavioral antecedents, as well as distinct aspects of the key concepts. In this way, it enables a valid formulation of the measurement items.

Conclusion

Currently, PBM research is incoherent, and the existing models conflict with one another. An in-depth understanding of the PBMs' differences is needed to facilitate their application in practice, as well as for their future development. In this study, we investigated whether the inconsistencies among the PBMs can be explained by their different background theories or their application domains. Our findings showed that the models' inconsistencies cannot be explained by the choice of their theoretical backgrounds. Instead, such inconsistencies seemed to follow from the disconnectedness of privacy behavior modeling. We found that the background theories are typically used for model construction in combination with other theories, and the PBMs are often developed in an eclectic way for a specific purpose, hence based on various assumptions. The application domain has often been considered in the model construction, either through the background theory selection or the incorporation of contextual factors into the PBM. However, the application domain does not account for the differences among the PBMs, and variations were also observed among the PBMs that were constructed for a similar domain. The PBMs seem to be derived from the *ad hoc*-type model contextualization, without a detailed analysis of the domain characteristics. Overall, because the current research has not yet been properly structured and seems to be evolving, the need and the possibilities for the development of more comprehensive and elaborate theories and model contextualization should be considered.

Funding

This work was supported by the Riitta ja Jorma J. Takanen Foundation, Tauno Tönning Research Foundation, and Jenny and Antti Wihuri Foundation.

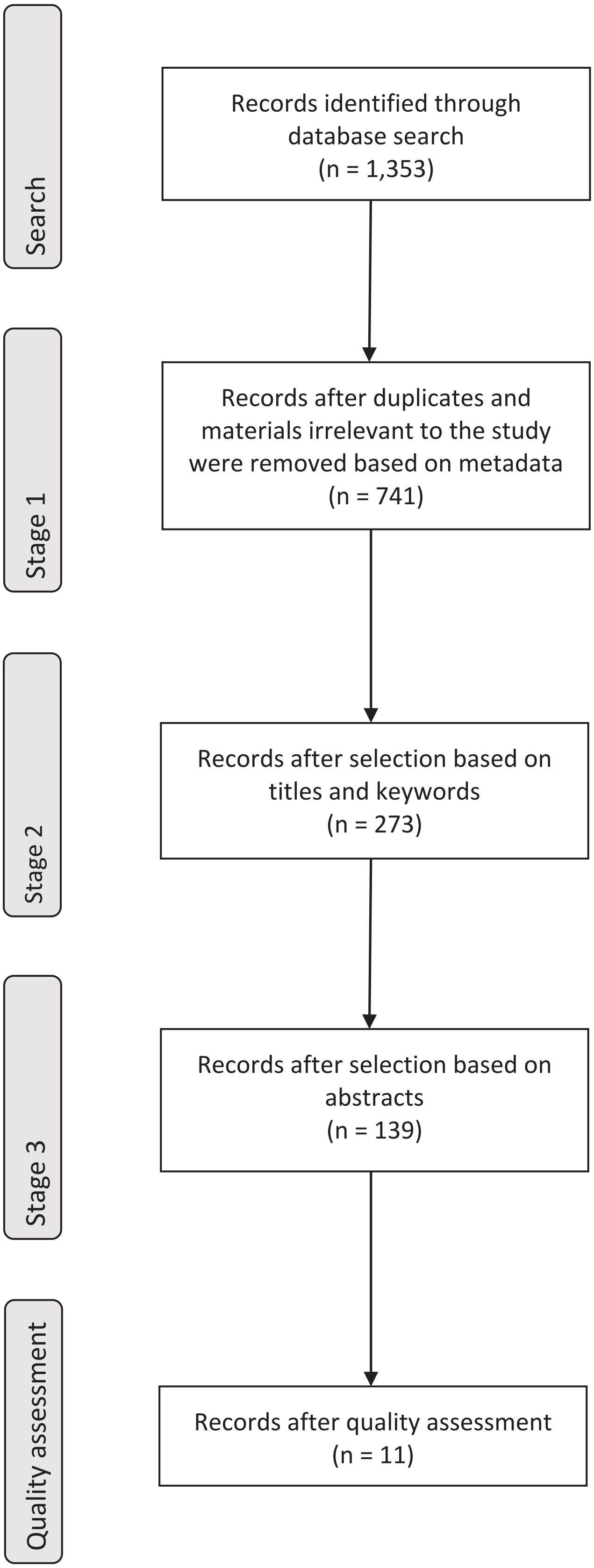
References

References preceded with an asterisk are included in the literature review.

1. Li Y. Empirical studies on online information privacy concerns: literature review and an integrative framework. Commun Assoc Inf Syst. 2011;28:453–96.
2. Smith HJ, Dinev T, Xu H. Information privacy research: an interdisciplinary review. MIS Quart. 2011;35:989–1015.
3. Culnan MJ, Armstrong PK. Information-privacy concerns, procedural fairness, and impersonal trust: an empirical investigation. Organ Sci. 1999;10:104–15.
4. *Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (IUIPC): the construct, the scale, and a causal model. Inf Syst Res. 2004;15:336–55.
5. *Dinev T, Hart P. An extended privacy-calculus model for ecommerce transactions. Inf Syst Res. 2006;17:61–80.
6. Van Slyke C, Shim JT, Johnson R, Jiang J. Concern for information privacy and online consumer purchasing. J Assoc Inf Syst. 2006;7:415–44.
7. Bansal G, Zahedi F, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern, and trust in disclosing health information online. Decis Supp Syst. 2010;49:138–50.
8. Chellappa RK, Sin RG. Personalization versus privacy: an empirical examination of the online consumer's dilemma. Inf Technol Manage. 2005;6:181–202.
9. *Kehr F, Kowatsch T, Wentzel D, Fleisch E. Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. Inf Syst J. 2015;25:607–35.
10. *Anderson CL, Agarwal R. The digitization of healthcare: boundary risks, emotion, and consumer willingness to disclose personal health information. Inf Syst Res. 2011;22:469–90.
11. Osatuyi B. Personality traits and information privacy concern on social media platforms. J Comput Inf Syst. 2015;55:11–19.
12. *Bansal G, Zahedi FM, Gefen D. Do context and personality matter? Trust and privacy concerns in disclosing private information online. Inf Manage. 2016;53:1–21.
13. Li Y. Theories in online information privacy research: a critical review and an integrated framework. Decis Supp Syst. 2012;54:471–81.
14. Dinev T, McConnell AR, Smith HJ. Research commentary – informing privacy research through information systems, psychology, and behavioral economics: thinking outside the "APCO" box. Inf Syst Res. 2015;26:639–55.
15. Kitchenham BA, Charters S. Guidelines for performing systematic literature reviews in software engineering. Staffordshire (UK): Keele University and Durham (UK): University of Durham; 2007. Technical report EBSE-2007-01.
16. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg. 2009;8(5):336–41.
17. Zhou Y, Zhang Y, Huang X, Yang S, Ali Babar M, Tang H. Quality assessment of systematic reviews in software engineering: a tertiary study. Paper presented at: EASE '15, 9th International Conference on Evaluation and Assessment in Software Engineering; 2015 Apr 27 –29; Nanjing, China.
18. Smith HJ, Milberg JS, Burke JS. Information privacy: measuring individuals' concerns about organizational practices. MIS Quart. 1996;20:167–96.
19. Stewart KA, Segars AH. An empirical examination of the concern for information privacy instrument. Inf Syst Res. 2002;13:36–49.
20. Castaneda JA, Montoso FJ, Luque T. The dimensionality of customer privacy concern on the Internet. Online Inf Rev. 2007;31:420–39.
21. *Mothersbaugh DL, Foxx WK, Beatty SE, Wang S. Disclosure antecedents in an online service context: the role of sensitivity of information. J Serv Res. 2012;15:76–98.
22. *Wakefield R. The influence of user affect in online information disclosure. J Strateg Inf Syst. 2013;22:157–74.
23. *Li Y. A multi-level model of individual information privacy beliefs. Electron Commer Res Appl. 2014;13:32–44.
24. *Lee H, Lim D, Kim H, Zo H, Ciganek AP. Compensation paradox: the influence of monetary rewards on user behaviour. Behav Inf Technol. 2015;34:45–56.
25. *Libaque-Saenz CF, Chang Y, Kim J, Park M-C, Rho JJ. The role of perceived information practices on consumers' intention to authorise secondary use of personal data. Behav Inf Technol. 2016;35:339–56.
26. *Ozturk B, Nusair K, Okumus F, Singh D. Understanding mobile hotel booking loyalty: an integration of privacy calculus theory and trust-risk framework. Inform Syst Front. 2017;19:753–67.
27. Campbell AJ. Relationship marketing in consumer markets: a comparison of managerial and consumer attitudes about information privacy. J Direct Marketing. 1997;11:44–57.
28. Donaldson T, Dunfee TW. Toward a unified conception of business ethics: integrative social contracts theory. Acad Manage Rev. 1994;19:252–84.
29. Son J-Y, Kim SS. Internet users' information privacy-protective responses: a taxonomy and a nomological model. MIS Quart. 2008;32:503–29.
30. Hoffman D, Novak T, Peralta M. Building consumer trust online. Commun ACM. 1999;42:80–85.
31. Xu H, Dinev T, Smith J, Hart P. Information privacy concerns: linking individual perceptions with institutional privacy assurances. J Assoc Inf Syst. 2011;12:798–824.
32. Li H, Sarathy R, Xu H. Understanding situational online information disclosure as a privacy calculus. J Comput Inf Syst. 2010;51:62–71.
33. Liao C, Liu CC, Chen K. Examining the impact of privacy, trust and risk perceptions beyond monetary transactions: an integrated model. Electron Commer Res Appl. 2011;10:702–15.
34. Li H, Sarathy R, Xu H. The role of affect and cognition on online consumers' decision to disclose personal information to unfamiliar online vendors. Decis Supp Syst. 2011;51:434–45.
35. Weible RJ. Privacy and data: an empirical study of the influence of types of data and situational context upon privacy perceptions [dissertation]. Starkville (MS): Mississippi State University; 1993.
36. Fishbein M, Ajzen I. Belief, attitude, intention, and behavior: An introduction to theory and research. Reading (MA): Addison-Wesley; 1975.
37. Ajzen I, Fishbein M. Understanding attitudes and predicting social behavior. Englewood Cliffs (NJ): Prentice-Hall; 1980.
38. Vroom VH. Work and motivation. New York (NY): Wiley; 1964.
39. Petronio S. Boundaries of privacy: Dialectics of disclosure. Albany (NY): SUNY Press; 2002.
40. Laufer RS, Wolfe M. Privacy as a concept and a social issue: a multidimensional development theory. J Soc Issues. 1997;33:22–42.
41. Acquisti A, Grossklags J. What can behavioral economics teach us about privacy? In: Acquisti A, De Capitani Di Vimercati S, Gritzalis S, Lambrinoudakis C, editors. Digital privacy: Theory, technologies and practices. Boca Raton (FL): Auerbach Publications (Taylor and Francis Group); 2007. p. 363–77.
42. Acquisti A, Brandimarte L, Loewenstein G. Privacy and human behavior in the age of information. Science. 2015;347:509–14.

43. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. Econometrica. 1979;47:263–91.
44. Tversky A, Kahneman D. Advances in prospect theory: cumulative representation of uncertainty. J Risk Uncertain. 1992;5:297–323.
45. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. Acad Manage Rev. 1995;20:709–34.
46. McKnight D, Cummings LL, Chervany NL. Initial trust formation in new organizational relationships. Acad Manage Rev. 1998;23:473–90.
47. Aven T, Renn O. On risk defined as an event where the outcome is uncertain. J Risk Res. 2009;12:1–11.
48. Kuhn TS. The structure of scientific revolutions. 4th ed. Chicago (IL): The University of Chicago Press; 1962.

Appendix A: Flow diagram of the literature review process


Search
Records identified through database search (n = 1,353)
Stage 1
Records after duplicates and materials irrelevant to the study were removed based on metadata (n = 741)
Stage 2
Records after selection based on titles and keywords (n = 273)
Stage 3
Records after selection based on abstracts (n = 139)
Quality assessment
Records after quality assessment (n = 11)

Appendix B: Inclusion and exclusion criteria of the study selection

Selection process	Inclusion criteria	Exclusion criteria
First round (metadata)	A journal article OR a conference paper OR a doctoral dissertation	Not written in English OR published in a forum irrelevant to PBM research OR a book chapter OR a technical report OR an opinion paper OR a presentation OR an interview OR a summary/extended abstract OR a master's thesis OR a duplicate
Second round (title and keywords)	AND presents an empirical model describing privacy behavior AND studies data subjects' aspect	AND clearly focuses on technical or legislative aspects of data protection
Third round (abstract)	AND presents a PBM with the behavioral outcome (willingness to disclose data OR adoption, usage, or intention to use an application or a service requiring personal data disclosure) AND presents an empirical study or PBM based on empirical results AND studies data subjects' (or users' or customers') aspect	AND the presented model does not focus on data subjects' point of view

Appendix C: Quality assessment checklist questions and their measurement scales

Question	Measurement scale
Reporting	
Is the study based on a previous model or a theory, and is it properly presented?	No/Partially/Yes/Not relevant
Are the study's objectives (such as constructing a new model or modifying an existing one) clearly defined?	No/Partially/Yes/Not relevant
Is the exact service or application context presented? (e.g., application type, if any, data types collected for the service or application, usage and processing of the collected data, data controller)	No/Partially/Yes/Not relevant
Are the study's end results with respect to its objectives really presented?	No/Partially/Yes/Not relevant
Rigor	
Are the data collection methods presented?	No/Partially/Yes/Not relevant
Are the collected data appropriate for constructing the type of statistical model presented in the paper? (e.g., Is the number of respondents large enough for SEM? Is the population representative? Are the variable types appropriate for the model in question?)	No/Partially/Yes/Not relevant
Are the measurements of the variables (i.e., question and scale) used in the study clearly defined?	No/Partially/Yes/Not relevant
Is the validity of the metrics discussed? (e.g., through Cronbach's alpha)	No/Partially/Yes/Not relevant
Are the characteristics of the data set presented? (e.g., distributions and means)	No/Partially/Yes/Not relevant
Are the quality aspects of the sample presented? (e.g., respondent types and demographic information, respondent recruitment strategy, response rate, sampling method)	No/Partially/Yes/Not relevant
Does the study provide descriptions of the data analysis methods?	No/Partially/Yes/Not relevant
Are the described data analysis methods appropriate for the purpose of the study?	No/Partially/Yes/Not relevant
Is the data analysis presented comprehensively? (e.g., *p*-values, R^2, factor loadings)	No/Partially/Yes/Not relevant
Credibility	
Does the study report clear, unambiguous findings based on evidence and arguments? (e.g., Are the findings logically derived from the data instead of presenting opinions, etc.?)	No/Partially/Yes/Not relevant
Does the study include a clear, comprehensive, and coherent validity discussion? (construct validity, internal validity, external validity, possibility of bias, limitations of the study)	No/Partially/Yes/Not relevant
Relevance	
Does the study's citation index prove it to be of high relevance?	Field-Weighted Citation Impact
Are the study's end results discussed with respect to those of previous studies?	No/Partially/Yes/Not relevant
Is the contribution of the study clearly presented?	No/Partially/Yes/Not relevant