Beyond Privacy Concerns: Examining Individual Interest in Privacy in the Machine Learning Era

Nicholas James Brown

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Business, Business Information Technology

Paul Benjamin Lowry, Chair
Quinton Nottingham, Co-Chair
Vitali Mindel
Dezhi Wu

May 8, 2023
Blacksburg, Virginia

Keywords: human involvement, data annotation, enhanced APCO, explainable AI, ML, privacy concerns, privacy interest, privacy protection behaviors, privacy paradox, privacy calculus, scale development

Beyond Privacy Concerns: Examining Individual Interest in Privacy in the Machine Learning Era

Nicholas James Brown

ABSTRACT

The deployment of human-augmented machine learning (ML) systems has become a recommended organizational best practice. ML systems use algorithms that rely on *training data* labeled by human annotators. However, human involvement in reviewing and labeling consumers' voice data to train speech recognition systems for Amazon Alexa, Microsoft Cortana, and the like has raised privacy concerns among consumers and privacy advocates. We use the enhanced APCO model as the theoretical lens to investigate how the disclosure of human involvement during the supervised machine learning process affects consumers' privacy decision making. In a scenario-based experiment with 499 participants, we present various company privacy policies to participants to examine their trust and privacy considerations, then ask them to share reasons why they would or would not opt in to share their voice data to train a companies' voice recognition software. We find that the perception of human involvement in the ML training process significantly influences participants' privacy-related concerns, which thereby mediate their decisions to share their voice data. Furthermore, we manipulate four factors of a privacy policy to operationalize various cognitive biases actively present in the minds of consumers and find that *default trust* and *salience biases* significantly affect participants' privacy decision making. Our results provide a deeper contextualized understanding of privacy-related concerns that may arise in human-augmented ML system configurations and highlight the managerial importance of considering the role of human involvement in supervised machine learning settings. Importantly, we introduce perceived human involvement as a new construct to the information privacy discourse.

Although ubiquitous data collection and increased privacy breaches have elevated the reported concerns of consumers, consumers' behaviors do not always match their stated privacy concerns. Researchers refer to this as the privacy paradox, and decades of information privacy research have identified a myriad of explanations why this paradox occurs. Yet the underlying crux of the explanations presumes privacy concern to be the appropriate proxy to measure privacy attitude and compare with actual privacy behavior. Often, privacy concerns are situational and can be elicited through the setup of boundary conditions and the framing of different privacy scenarios. Drawing on the cognitive model of empowerment and interest, we propose a multidimensional privacy interest construct that captures consumers' situational and dispositional attitudes toward privacy, which can serve as a more robust measure in conditions leading to the privacy paradox. We define privacy interest as a consumer's general feeling toward reengaging particular behaviors that increase their information privacy. This construct comprises four dimensions—impact, awareness, meaningfulness, and competence—and is conceptualized as a consumer's assessment of contextual factors affecting their privacy perceptions and their global predisposition to respond to those factors. Importantly, interest was originally included in the privacy calculus but is largely absent in privacy studies and theoretical conceptualizations. Following MacKenzie et al. (2011), we developed and empirically validated a privacy interest scale. This study contributes to privacy research and practice

by reconceptualizing a construct in the original privacy calculus theory and offering a renewed theoretical lens through which to view consumers' privacy attitudes and behaviors.

Beyond Privacy Concerns: Examining Individual Interest in Privacy in the Machine Learning Era

Nicholas James Brown

GENERAL AUDIENCE ABSTRACT

The deployment of human-augmented machine learning (ML) systems has become a recommended organizational best practice. ML systems use algorithms that rely on training data labeled by human annotators. However, human involvement in reviewing and labeling consumers' voice data to train speech recognition systems for Amazon Alexa, Microsoft Cortana, and the like has raised privacy concerns among consumers and privacy advocates. We investigate how the disclosure of human involvement during the supervised machine learning process affects consumers' privacy decision making and find that the perception of human involvement in the ML training process significantly influences participants' privacy-related concerns. This thereby influences their decisions to share their voice data. Our results highlight the importance of understanding consumers' willingness to contribute their data to generate complete and diverse data sets to help companies reduce algorithmic biases and systematic unfairness in the decisions and outputs rendered by ML systems.

Although ubiquitous data collection and increased privacy breaches have elevated the reported concerns of consumers, consumers' behaviors do not always match their stated privacy concerns. This is referred to as the privacy paradox, and decades of information privacy research have identified a myriad of explanations why this paradox occurs. Yet the underlying crux of the explanations presumes privacy concern to be the appropriate proxy to measure privacy attitude and compare with actual privacy behavior. We propose privacy interest as an alternative to privacy concern and assert that it can serve as a more robust measure in conditions leading to the privacy paradox. We define privacy interest as a consumer's general feeling toward reengaging particular behaviors that increase their information privacy. We found that privacy interest was more effective than privacy concern in predicting consumers' mobilization behaviors, such as publicly complaining about privacy issues to companies and third-party organizations, requesting to remove their information from company databases, and reducing their self-disclosure behaviors. By contrast, privacy concern was more effective than privacy interest in predicting consumers' behaviors to misrepresent their identity. By developing and empirically validating the privacy interest scale, we offer interest in privacy as a renewed theoretical lens through which to view consumers' privacy attitudes and behaviors.

ACKNOWLEDGEMENTS

This dissertation is the culmination of eight long years in graduate school, where I received high-quality instruction, advisement, and encouragement from some of the most intelligent, most influential people I have ever met. I started my academic journey with the intention to obtain one graduate degree and ended up with five from two universities. Part of the reason for my prolonged stay in the halls of Virginia Tech is that I admired my professors and wanted desperately to emulate them, their demeanor, their kindness, and more. Writing this dissertation is my first step in treading the path paved by my predecessors who graciously illuminated their walk so that I could follow. With giant shoes to fill, I stumbled often trying to follow in their shadows. Every time, they helped me to regain my footing, providing valuable tips on how I should move forward. Eight years later, I have compiled a vast list of remarkable people to whom I am deeply indebted for helping me to the finish line. Thanks to these individuals not only did I finish the race, but I also developed an endurance that will help me through the marathon ahead. What started as a crawl is now a steady jog. While everyone's names and acts of kindness race through my mind, I am terrified of committing any sins of omission. If this happens, please know that you will always have a special place in my heart—the memories we shared and the contributions you have made to my development are with me always.

Sitting at the apex of my "thank you" list is my dissertation committee. To Paul Lowry and Quinton Nottingham, my advisors. Paul, you have carefully crafted me into a junior scholar, spending copious hours, days, months, and years mentoring me in every facet of academic life from theory building with paradoxes and mysteries, to investigating research discourses for tensions, to edifying peers through gem-polishing reviews, to conducting rigorous scientific research, in general. Rarely do people have the honor to learn from a generational figure like you. I am deeply honored to have shared my journey with you and to belong to a lineage of extraordinary researchers who have also shared in the wondrous experience to learn from an information systems hegemony. My first research seminar was with you, and I am overjoyed that our research interests intersected. I dare not imagine who I would be as a researcher without you. The opportunity to work together is indelibly impressed in my heart and mind. Who I am now is thanks in large part to you and your dedication to training doctoral students.

Quinton, to say that I would not be here without you is an understatement. I viewed you as an instrumental role model to me from day one when I had the privilege to serve as your graduate assistant. Through our many conversations ranging from research to life, you always imparted pithy words of wisdom that immediately influenced my perspective on things. Many changes I have made in my life from only a few syllables you spoke to me. I remember hanging out at your home with Phil Thompson (shout out to Phil, too!) and reaching full enlightenment through a conversation profoundly reverberating in my mind still today. That was the day I learned the intricacies of life as a doctoral student and as a professor. It was as if someone provided me with a crystal ball into the future, but the brilliant images envisaged were too fluorescent for my eyes to see. However, you saw everything clearly. Your clairvoyance was indomitable—what you said came true. Talk about manifesting visions. I am honored to have received your encouragement, support, and continued vote of confidence in me. Your faith in me went a long way—sprinkle in lessons on mental fortitude, stick-to-itiveness, and perspicacity, and you've taught me the

essence of a "championship" mindset. And for this and many more reasons, I will be forever grateful to you.

To Vitali. I remember when you stopped by theory-building class and graciously accepted my invitation to meet and chat. I thought I was a "cool" instructor until I had the chance to converse with you. You may be pound for pound the coolest professor in academia. I hope to exude the same charm, confidence, and bravado when I become an assistant professor. You are equally brain and brawn, as you hit the weights as much as you hit those elites. I typically think in terms of "either/or"; however, seeing how you can balance success with personal wellness, theory with empirical, I can envision a future where "both" becomes my way of thinking. I admire your ability to build theories and design empirical studies. You are a standalone team because you are equipped with a breadth of skills and diversity of experiences. As much as I try to be anticipatory, I always learn new perspectives on various phenomena after conversing with you. You can sniff things that others cannot. When an interesting idea is trapped in a smorgasbord of words, you can detect its presence, extract it, and cook it into something appealing for others to delight in. If anyone can develop "revolutionary" from the "evolutionary," it would be you, and I am grateful to have witnessed your genius. You have become a great motivation to me as someone I will model in the early stage of my professorship career.

To Dezhi. The saying, "where the rubber meets the road," is a fitting description of how our collaborations materialized into quick (and successful) outcomes together. I learned from you the importance of execution. Ideating could only go for so long, but when inspiration struck, you were the first to help me light it ablaze. No deadline was too tight for a couple of great minds to get together, develop a plan of action, and execute it with precision. Working on your co-authorship teams granted me privileged insights into the manuscript writing and review process unparalleled by any scholastic training I had received. Working with you accelerated my learning and growth as a researcher, and I express my many thanks for our weekly meetings, systems testing, brainstorming, and many more cherished times together that contributed to my knowledge base as a researcher, student, and person. Your high level of care and compassion was greatly appreciated, too, as it helped fuel me throughout my doctoral journey. The journey was made all the better with you by my side.

Next, it is imperative that I acknowledge the instrumental figures who prepared me for my doctoral journey. Their influence goes beyond what they may have imagined. To them, I may have been one of many students in the world who matriculated into their master's program; but to me, they meant the world to me for pointing me northward toward academia. I thank them for their keen role in my formation as a burgeoning researcher and instructor. Frankly, I would like to thank them for being wholehearted friends to me.

To Parviz Ghandforoush and Lara Khansa, my early mentors. Parviz, truly, you put me in a position of success. Our conversations in your office about course design, course delivery, syllabus writing, student development, and more were foundational to creating my teaching identity and helping me to develop rapport with our prominent MBA and MIT students. Through your expert wisdom in leading faculty and students, I learned the characteristics of a transformational leader and the importance of applying an ethic of care in all our correspondence. I am fortunate to have you as a mentor in my life.

Lara, I have spent the majority of my active academic career with you by my side. You have always been my evangelist, expressing your confidence in me to top leadership at

Virginia Tech. The blessings you have provided to me are innumerable, and I could only hope to pass along just a fraction of those blessings to you that you provided to me. Serving as your distance learning instructor, developing a database course together, and conversing with you in the hallways of Pamplin or on the phone represent only a sliver of the joyful activities I had the privilege to share with you.

I would like to send an especial thanks to the professors who trained me during my master's training in the MBA, Master's of IT, and Statistics departments, and during my doctoral training in the BIT department. My intellectual development was sharpened by the courses, seminars, and mentorship I received from scholars in a myriad of disciplines. It was the erudition exhibited by these scholars that inspired me to become an academician. I hope to be as inspirational to others as these scholars were to me. The gratitude I extend to them for their dedication to high-quality instruction cannot be fully expressed in this section. I thank God for placing these individuals in my life. All I can wish is that the benevolence they bestowed on me is repaid tenfold in the future by friends, family, students, and others—an endless virtuous circle:

Idris Adjerid (for writing the seminal 2018 MISQ Paper of the Year that I revisit annually as inspiration to my research); Ralph Badinelli (for exposing me to Internet of Things and smart services); Sheryl Ball (for teaching me the difference between economics and psychology experiments); Daniel Beal (for the exceptional training on survey instrument development); Sudip Bhattacharjee (for never letting me forget the basics of T-accounts, credits, and debits); James Campbell (for the keen insights and instructive labs on remote sensing); Kevin Carlson (for allowing me to serve on the Online Learning Task Force); Laurence Carstensen (for teaching me how to create ArcGIS web maps); Dipankar Chakravarti (for the foundational training on qualitative research); Anne Driscoll (for encouraging me to join the Stats Department, which has been a godsend by the way); Joseph Gabbard (for introducing me to human cognition and processing theories, many of which I use in my research); Susan Gates (for facilitating rich conversations on ethics among students in our blog site); Kendall Giles (for arguably the best cybersecurity course I have taken at Virginia Tech); Feng Guo (for teaching me how to perform PCAs and factor analyses); Donald Hatfield (for hosting the yearly case competitions and poster sessions and for allowing me to serve as a judge); Dave Higdon (for guiding me through the Data Analysis and Stats Master's program as my chair and mentor); Barbara Hoopes (for taking me to Scandinavia study abroad and for the extremely helpful advice before I interviewed with various doctoral programs); Tabitha James (for your dedication to build up doctoral students in our weekly colloquium); Mahmood Khan (for taking me to India and Dubai study abroad and for the witty sense of humor that still makes me chuckle when I think of the good times together); Thomas Koch (for giving me a great sense of accomplishment when I passed your three econ exams); Greg Kulczycki (for teaching me how to program Java apps, build SQL databases, and develop web apps); Hamdy Mahmoud (for teaching me everything I know about design of experiments); Raymond Major (for being an inspiration to me to pursue academia and to become skilled in statistics); Sattar Mansi (for teaching me to keep an eye on the 10-year treasury in anything I do financially); David McPherson (for introducing me to the field of electrical engineering); Naren Ramakrishnan (for showing me the power of social media analytics before I even learned to use social media); Robert Settlage (for the exceptional training in R and R Studio); Tom Sheehan (everything I know about web apps is thanks to you); Steven Sheetz (I took a majority of

your database courses because I enjoyed them so much); Matt Slifko (for sitting with me to teach the basics of linear algebra in your advanced regressions course); Jennifer Van Mullekom (for allowing me to serve as a statistics consultant in SAIG); Viswanath Venkatesh (for sharing your profound wisdom in your Road to Success and Mixed Methods books); Alan Wang (for exposing me to design science research); and Andrew Watson (for introducing me to management theories that eventually became a stark fascination of mine).

Of all the people at Virginia Tech, I must also thank the program directors who gave me the chance to enter into their respective programs. First, and foremost, to Dana Hansson, it sounds cliché, but "without you, there would be no me." I remember the fateful day in 2015 when you called to personally inform me that Virginia Tech had accepted me into its MBA program. I was in the middle of a workout when I received your call. When you introduced yourself to me, I immediately stepped out of the gym to listen to every word you would say. When you informed me of the terrific news, I felt so overjoyed that I immediately left to tell my parents. The next day, I realized I had left my wallet and the rest of my belongings at the gym. Talk about a euphoric trance you set off in me with only one word, "accepted."

To Terry Hinders. Working with you in the MIT program opened my eyes to the world of technology. I was once blind, and our conversations allowed me to see. Seeing the high-caliber students in your program prompted me to jump on the bandwagon, and I am so happy I did. So much so that I recommended many of my MBA peers to join the MIT program, and vice versa, I recommended my MIT peers to join the MBA. Your scope of influence transcended beyond me to a dozen others. Thank you for being a powerful leader.

To Cliff Ragsdale. When I met you as a master's student, you instantly saw the potential in me to succeed at the doctoral level. I turned around to check whether I had wings poking out of my back because you predicted that I would soar with the eagles. Now, it took some time for me to spread those wings, but when the wind blew my way, I happened to take flight at the right time. Thank you for bringing me into the BIT Ph.D. program. The special place I hold in my heart for having you in my life is irreplaceable.

To Robin Russell. You always had students' best interest at heart and were invested in our success. You were the first to acknowledge our noteworthy accomplishments and the first to lend support—financially, academically, and emotionally. Through personal phone calls, Zoom meetings, and text messages, you made yourself fully available to us, demonstrating your touch of empathy and your philosophy to never leave any student behind. I was fortunate to enter into the BIT program, but I was more fortunate to have you as a leader and mentor. Thank you for wearing many hats and for gracefully switching any one of them at a moment's notice.

To the staff of the MBA and MIT programs and the BIT and Stats Departments, your warm presence permeated the hallways in Northern Virginia and Blacksburg campuses and made the office a welcoming place. The atmosphere was cordial, the laughter was blissful, and the operation was rolling thanks to you. The NCAA March Madness tourneys were a delight as well. I'm glad I was able to hang my banner among the greats. Beverly Griffin, Holly Gillcash, Christine Aquino, Kathy Orton, Tracy McCoy, Teena Long, Christina Smith, Robin Littleton, Christina Dillon, Adrienne Sable, and Shea Walters—my many thanks to you for making my time at Virginia Tech very special.

In addition, many thanks to my peers in the Graduate Student Assembly at the NCR and Blacksburg campuses. Especial thanks to our Graduate School deans, Karen DePauw and Aimee Surprenant, for their exceptional leadership. If any student wishes to pursue graduate education, one reason to come to Virginia Tech is for our Graduate School, the orchestrator of instant rapport building among a network of like-minded grad students. The saying, "This is Home" comes to life because of our Grad School.

Speaking of "home," the BIT Department was my second home—virtually and physically. Prior to the pandemic, I stayed on campus from 8am to 8pm, Monday through Friday, attending classes, working on GA assignments, and writing papers. After the pandemic, nothing changed except my work hours. Working virtually allowed me to start my days sooner and end them later—but instead for seven days a week, yikes. I would like to thank all of the faculty for providing a safe place to converse, study, and relax. Thank you for the uplifting edification: Alan Abrahams, Ruba Aljafari, Mike Gordon, Alice Jang, Jiayi Liu, Michelle Seref, Onur Seref, Wenqi Shen, David Simpson, Tony Vance, and Chris Zobel.

Shout out to the PhD Project for their endless support and encouragement. Virginia Tech was my home, but the PhD Project was my family. Every person I met in the ISDSA, Annual Conference, symposia, and more unabashedly treated me like a brother: "Nick, whatever you need, I got you!" The Project is doing God's work because the organization is changing lives. If anyone in the Project runs for a government office, they will have my vote along with the thousands I will strum up to get them elected. My most sincere thanks to Blane Ruschak, Bernie Milano, Myrna Varner, Marie Zara, Tara Perino, Cristina Pazos, and the entire ISDSA faculty and students. My life is more abundant with you in it. I have many special stories to tell from my terrific experience with Project faculty; I will be sure to tell you personally when I see you because now that I am family, you cannot get rid of me, hehe. I'll see you at many more special occasion events in the future.

My classmates at Virginia Tech were the best peers a person could ever ask for: supportive, encouraging, ambitious, and successful. My motivational drive to do well scholastically is due in part to them and my trying to keep up. I very much appreciate our time together, the camaraderie, and for making every day enjoyable when learning in class or hanging out in the office together. Every interaction I had was meaningful, which makes it nearly impossible to enumerate every peer who contributed to my scholastic success and personal development. If I forget to name you, let's blame it on being too occupied with writing this dissertation: Carolina, Anthony, Rachel, Kate, Sobe, Samantha, multiple Chris's, Mahsa, Awide, Patrice, multiple Steve's, Kathleen, multiple John's, Pratik, Chloe, Brandon, Justin, Natalie, Chaimaa, Peter, Sagar, Ipek, Duygu, Cam, Varada, George, Jeremy, Dan, David, Zeynep, Shaokang, Mohammad, Behnam, Frank, Jim, Kira, Autumn, Ariane, Barrett, Jessica, Leonard, Taylor, Bryon, Kevin, Joey, Swapna, Kara, Paul, Jonathan, Paula, Elizabeth, Hannah, Farida, Jackie, Brent, Kelly, Darryl, and so many more. It was a fun run, whether it was hopping time zones together or a friendly comment in class, indelible memories were created.

Finally, to my family: George, Ki, Sinclair, Ben, you have been there for me since day one. Your patience, understanding, cheers, and support were clutch—I could not have made a nice comeback without you. I share the spoils of this victory all with you.

Table of Contents

Chapter 1: The Privacy-Explainability Paradox in Machine Learning Systems—An Empirical Examination of the Role of Human Involvement in Training Data Privacy

1.1. Introduction

Machine learning (ML) systems differ from traditional information systems (IS) (e.g., expert systems, intelligent agents, decision support systems), such that companies cannot readily explain the functionality, decision making, and outputs of an ML system by viewing the source code of the algorithm (Grant & Wischik, 2020; Villaronga et al., 2018). Employees must perform a careful review of curated training datasets to explain which types of data features are inferred by the system and why (Grant & Wischik, 2020). However, the review of consumers' data by human agents infringes on consumers' rights to privacy and their ability to control their information privacy (Bélanger & Crossler, 2011; Pavlou, 2011; Smith et al., 2011).

Conversational ML systems (also referred to as voice agents) interface directly with consumers and use induction to draw inferences from trained datasets, whereby new stimuli, such as newly spoken voice commands, are algorithmically matched with labeled data, and responses are retrieved and rendered from predetermined classifications of output (Benbya et al., 2021; Grant & Wischik, 2020; Lebovitz et al., 2021). These ML systems extract and analyze billions of data points, including sensitive, specific, and identifiable information (Sutanto et al., 2013) on individual consumers to create "super profiles" (Al-Natour et al., 2020). Data from super profiles are then used to compute knowledge in the form of parameter weights and strengths of connections in a neural network to seemingly learn consumers' behaviors and tendencies (Zuboff, 2015, 2019). Consequently, data privacy regulations have been enacted to grant consumers who have given control of their data to companies specific rights regarding the collection, handling, and use of their data (General Assembly of Virginia, 2021; Office of the Attorney General of California, 2021; Parliament and Council of the European Union, 2016). Meanwhile, researchers have issued

calls for greater transparency and explainability of the actions of ML algorithms to evaluate the fairness and biases inherent in these systems (e.g., Rai, 2020; Rai et al., 2019; Teodorescu et al., 2021).

Recent research, however, reveals a dichotomy between privacy and explainability, a tension called the *privacy-explainability paradox* (Grant & Wischik, 2020), inherent in ML systems and the privacy policies to which consumers consent. Explaining the actions of voice systems requires a review of the personal voice data from which the voice agents are trained. This allows human agents to evaluate and train the data for accuracy, prevent the perpetuation of biases, and furnish explanations for ML systems outputs. Companies use privacy policies as vehicles to transparently explain this manual review practice but can exploit consumers' cognitive biases such as *default trust bias*, *salience bias*, or *heuristic processing* of information to obtain consent to do so (Dinev et al., 2015; Gerlach et al., 2019). We find that extant empirical studies have not investigated the mechanisms through which perceived human involvement influences consumers' privacy decision making in the context of voice system use (e.g., Amazon Alexa, Alibaba AliGenie, Apple Siri, Google Home). Thus, understanding the mechanisms that explain how *human involvement* in machine learning practices may factor into consumers' privacy decisions is crucial because such understanding will enable companies to design interventions to better protect consumers' privacy and promote their willingness to share voice data to train companies' voice recognition algorithms.

We posit that human involvement in conversational ML systems can serve as an important antecedent to privacy considerations and can reduce consumers' willingness to share their training data. This is highly undesirable if companies require balanced, diverse training datasets to prevent biases and discrimination in the training of their ML. To investigate human involvement, we build on the human-ML augmentation research related to big data, AI, and biases in ML systems (Kane

et al., 2021; Teodorescu et al., 2021) and use the lens of nondata breach privacy concerns (Brooks et al., 2017) to explain how consumers can have privacy concerns outside of data theft or security breach events (Kaufman et al., 2009). Although research shows the extensive benefits of the conjoined agency between humans and ML systems (Murray et al., 2020; Raisch & Krakowski, 2021) and human-in-the-loop system configurations (Grønsund & Aanestad, 2020), oftentimes the research contexts are within organizations, and the benefits to achieve fairness in ML decisions (Teodorescu et al., 2021) or enhance explainability (Rai et al., 2019) are reasons why human-ML augmentation is preferred. The role of human involvement in enhancing machine learning outcomes for organizations and societies is a key factor, but its comprehension at an individual level remains limited. Specifically, there is a need to explore how consumers perceive human involvement in this process and whether it influences their decisions regarding personal information disclosure and data sharing. Researchers and practitioners thus lack theoretical and empirical understanding on the effect of the knowledge of human involvement in voice data annotation practices on consumers' trust and privacy considerations. Consequently, our research objective is to investigate and explain the role of human involvement in the context of training data privacy. We believe that answering these questions can lead to novel, important contributions to the human-ML augmentation research discourse.

RQ1. How does the perception of human involvement in the training of conversational ML systems affect users' willingness to share their voice data with a company?

RQ2. What are the effects of privacy policy context and cognitive biases on users' privacy decision-making outcomes?

To answer these research questions, we ground our research model on the theoretical tenets of the enhanced antecedents-privacy concerns-outcomes (APCO) model because, in addition to trust

and privacy calculus, it theorizes effects of cognitive biases and information processing, both of which are prominent in privacy-decision making situations (Acquisti et al., 2013; Adjerid et al., 2018b). We operationalize privacy policies as a company's signal to increase its transparency toward its data processing and employ a scenario-based quantitative experiment using a between-subjects design and post hoc qualitative analysis to investigate the role of human involvement in the processing of voice training data. We provide quantitative empirical evidence and qualitative confirmation that perceived human involvement in the training of ML systems is viewed negatively by consumers, such that privacy concern and privacy risk are significantly influenced by perceived human involvement and mediate consumers' willingness to share voice data to improve companies' products and services.

1.2. Theoretical Development

We begin by tracing the roots of IS scholarship to ML systems, then explain the roles of human agency in both IS and ML systems. IS research on AI began in the 1970s when computers became integral systems embedded into workplaces to automate business practices and electronically process transactions (Benbya et al., 2021). First, *decision support systems* were the earliest development and relied on extensive databases from which decision makers could query to generate reports and interpret aggregated data (Alter, 1978; Turban & Watkins, 1986).

Second, *expert systems* were designed to draw inferences from a knowledge base and to accompany explanations with the recommendations they provided (Turban & Watkins, 1986). However, expert systems could not do this without human agency.

Third, *intelligent agents* functioned autonomously by using structured data input from relational databases but could not process unstructured data like video, audio, and images (Schuetz & Venkatesh, 2020). They also required human programmers to transform the unstructured data into usable formats (Benbya et al., 2021).

Finally, *cognitive computing systems* (Schuetz & Venkatesh, 2020) or *agentic IS artifacts* (Baird & Maruping, 2021) incorporated artificial intelligence to enable humanlike perceptual capabilities to sense and interpret unstructured stimuli in their environments—but humans agents remain prominently involved in the *training*, *explaining*, and *reviewing* of the ML algorithms (Ågerfalk, 2020; Asatiani et al., 2021; De Cremer & Kasparov, 2021; Lebovitz et al., 2021; Murray et al., 2020; Rai et al., 2019; Raisch & Krakowski, 2021) (see Appendix 1A). In summary, traditional IS depend on *human agents* to establish rules and define problem sets in which the systems can operate (Borges et al., 2021), whereas agentic IS artifacts can serve as *substitute agents* to humans in behavior- and outcome-based decision-making contexts (Baird & Maruping, 2021).

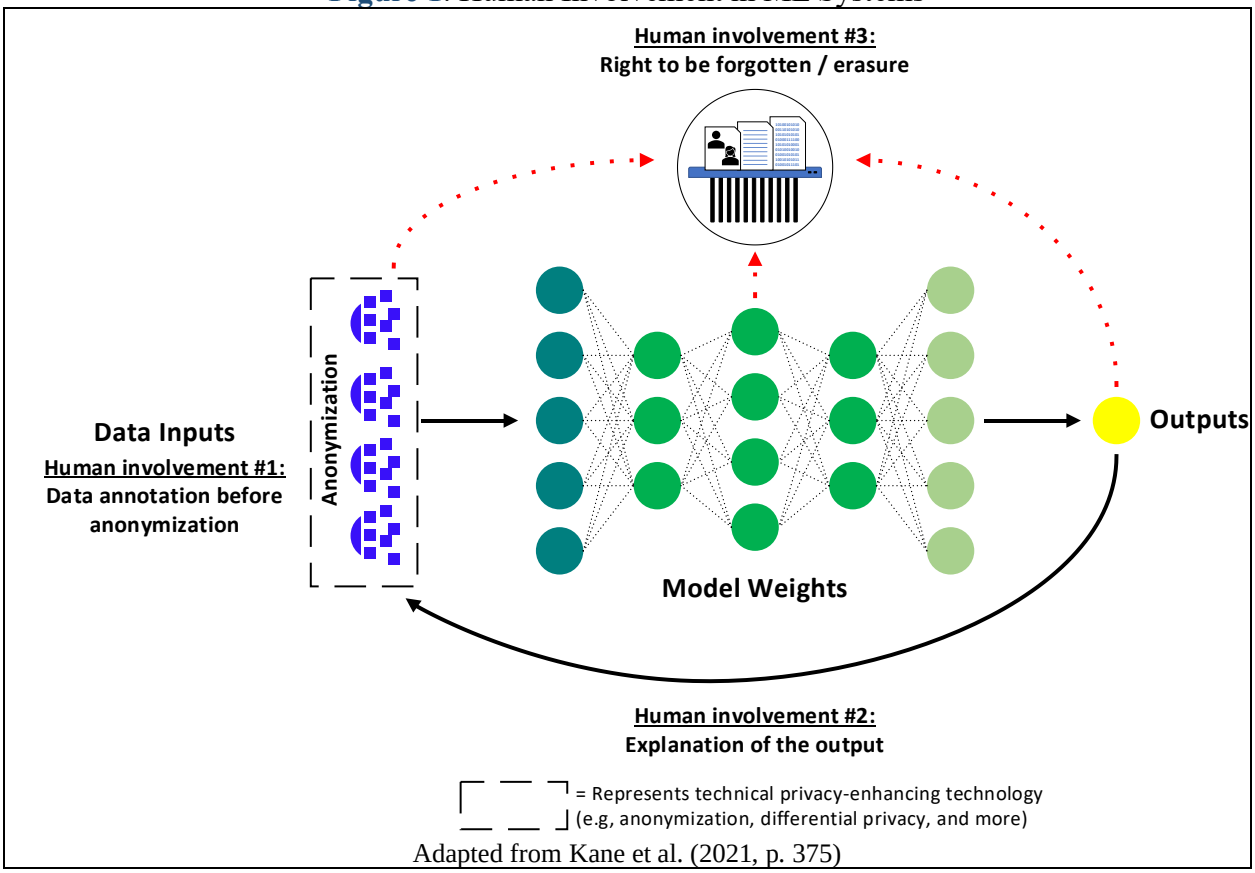
1.2.1. Human Involvement in ML Systems

Human involvement in ML systems is strongly recommended as an organizational best practice because of the sociotechnological context in which ML systems are deployed (Ågerfalk, 2020; Fügener et al., 2021; Grønsund & Aanestad, 2020). Configurations such as humans-in-the-loop and human-centered artificial intelligence are common, and c*onjoined agency* is an emerging concept in the research discourse that explains the types of integrated agencies between human agents and agentic IS agents (Baird & Maruping, 2021; Murray et al., 2020). Whereas a*gentic IS artifacts* possess the capacity "to constrain, complement, and substitute for humans in the practice of routines," (Murray et al., 2020, p. 553), automated technologies can make incomplete or flawed decisions based on spurious correlations found in the training data (Murray et al., 2020). Consequently, the boundaries of ML systems are "still managed by humans within technological, organizational, and institutional frames" (Ågerfalk, 2020, p. 5), primarily in the contexts of (1) training, (2) explaining, and (3) reviewing of data for compliance purposes (**Figure 1**).

First, human agents are needed during the data annotation stage to label unstructured data in

training datasets. Arguably, the availability of training datasets is the most important breakthrough catalyzing the surge in ML systems development and use (Wissner-Gross, 2016). This is because researchers and developers focused on advancing algorithms through computing power and increased storage to process complex algorithms (Benbya et al., 2021; Schuetz & Venkatesh, 2020), yet it was only until the availability of rich and high-quality datasets in the 2000s that breakthrough advancements in machine learning accelerated (Wissner-Gross, 2016).

Figure 1. Human Involvement in ML Systems



Adapted from Kane et al. (2021, p. 375)

Specifically, the combination of sensors, cyber-physical systems, and the internet enables the collection of vast amounts of audio-visual data. This data, in turn, allows human agents to swiftly train machines with remarkable computational power to recognize, group, and categorize human sounds and visual cues within a social context (Ågerfalk, 2020; Baird & Maruping, 2021; Benbya

et al., 2021; Kane et al., 2021; Lindebaum et al., 2019; Schuetz & Venkatesh, 2020). Thus, the perceptual capabilities for machines to *see* through computer vision and *communicate verbally* through speech recognition exponentially improved during the past two decades due to the swelling availability of consumer data to train ML models (Benbya et al., 2021; Kane et al., 2021).

Second, human agents are needed to interpret and explain the outputs from an ML algorithm. An inherent characteristic of ML systems is their functional opacity (Schuetz & Venkatesh, 2020). To increase the transparency of ML systems, human involvement is needed to remove the nonlinearity and unpredictability of algorithms' actions (Lindebaum et al., 2019). Rai (2020) refers to this as the *glass-box* model. Human agents review and explain algorithmic outputs to ensure that the outcomes are interpretable to human users. For example, in the case of conversational ML, human agents must discern between random noise, animal sounds, and human voices to ensure the speech detection patterns are correct. Ensuring that smart devices are not unintentionally activated by random noises is crucial in preventing accidental recordings (Merrill, 2020).

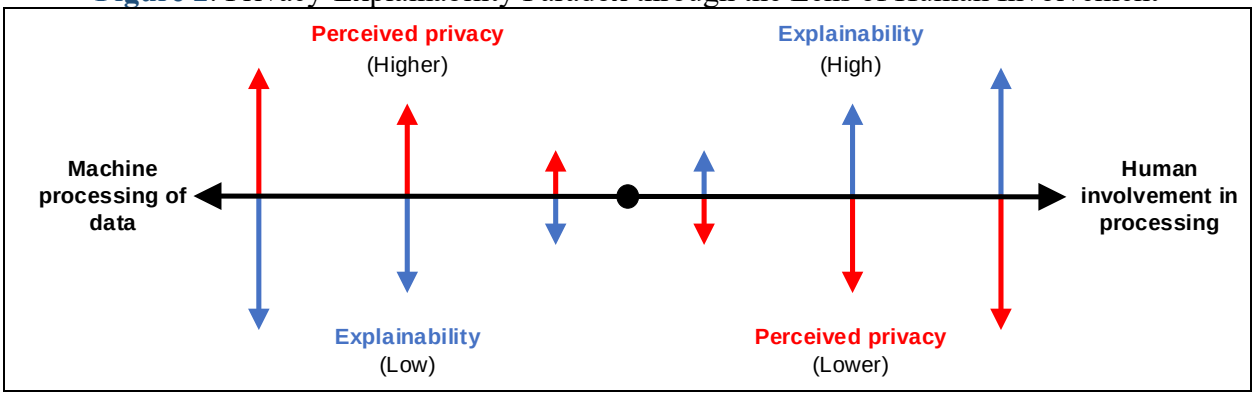
Last, human agents must sometimes review consumers' data for regulatory compliance reasons to process "right to be forgotten" and data erasure requests. For example, to request personal data erasure at Google, a form must be completed and human reviewers will consider how sensitive or private the content is before rendering a decision to erase personal data (Google, 2021). If the right to be forgotten request is granted, then the consumer's data must also be erased from the ML model weights and the training datasets. Computer scientists refer to this process as "machine unlearning" (Bourtoule et al., 2021), and scholars argue that it is nearly impossible to make ML systems forget, because of the impracticality to interpret model weights to ensure complete erasure of consumers' data (Villaronga et al., 2018).

1.2.2. Privacy-Explainability Paradox

Human involvement leads to a *privacy-explainability paradox* because of two explicit stipulations

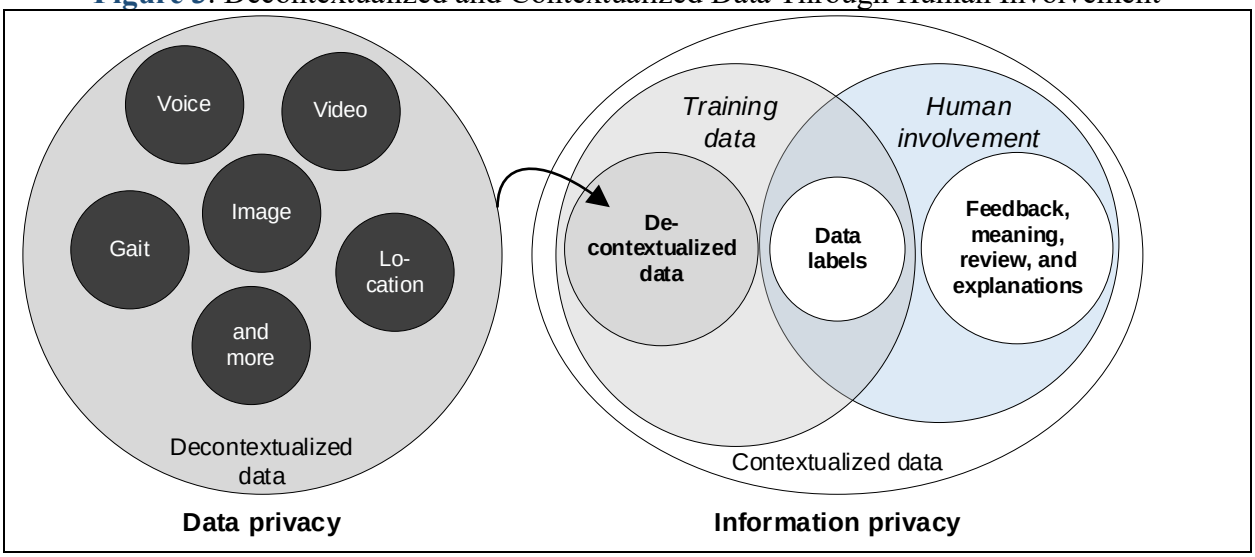
in comprehensive data privacy laws that are in direct conflict: (1) *individual's right to privacy* and (2) *individual's right to receive an explanation* on algorithmically derived outcomes—otherwise known as *explainability* (Grant & Wischik, 2020). The tension occurs because to explain how ML systems derive their logic, human agents must examine the data by which the algorithms are trained. **Figure 2** visualizes the privacy-explainability paradox as a continuum through the lens of human involvement.

Figure 2. Privacy-Explainability Paradox through the Lens of Human Involvement



This continuum illustrates that when systems alone process consumers' data, then perceived information privacy is higher but explainability of an algorithmic outcome is low. Conversely, when human agents process consumers' data, they can ascribe contextual meaning to the abstract data, which can infringe on information privacy rights but enhance the explainability of an algorithmic outcome. Ågerfalk (2020) refers to raw data as *decontextualized* data that do not contain any socially meaningful signs. Particularly, humans do the sensemaking of the emergent features detected in pattern recognition, and they can interpret the contextual relevance of the data and consider various uses and explanations of the data in practice (Ågerfalk, 2020). This practice of ground truth labeling shifts data privacy to information privacy, a notable issue in IS (**Figure 3**).

Figure 3. Decontextualized and Contextualized Data Through Human Involvement



1.2.3. Enhanced Antecedents-Privacy Concerns-Outcomes Model

Because the concept of privacy is multidimensional and is viewed differently among the social sciences (Laufer & Wolfe, 1977), *privacy concern* is operationalized as a proxy for privacy (Malhotra et al., 2004). *Privacy concern* is the belief that a person's privacy is at risk of being violated (Culnan & Armstrong, 1999). The broader *information privacy* is defined as the rights of a person to determine to what extent their personal information is shared and when and how it is shared with others (Westin, 1967). Dinev et al. (2015) proposed the *enhanced* APCO model that subsumed three privacy concern macromodels that at the time "summarized almost all of the positivist empirical assessments of privacy up to that date" (Dinev et al., 2015, p. 639). The original APCO model includes the constructs of information privacy concern, privacy risk, trust, and benefits (Smith et al., 2011). *Information privacy concern* comprises three dimensions: perceived surveillance, perceived intrusion, and secondary use of personal information (Xu et al., 2012a). *Privacy risks* are perceptions that a potential loss of data or misuse may occur when releasing personal information to an organization (Dowling & Staelin, 1994). *Trust* toward the data controller is the degree to which consumers believe the controller will reliably safeguard their

collected personal information (Gefen et al., 2003). Specifically, trust beliefs comprise three dimensions: *competence* (ability of the company to do what the consumer needs), *benevolence* (belief that the company acts in the consumers' best interest), and *integrity* (belief that the company is honest) (McKnight et al., 2002; McKnight et al., 1998). *Perceived benefits* are the net positive outcomes of disclosing personal information (Smith et al., 2011).

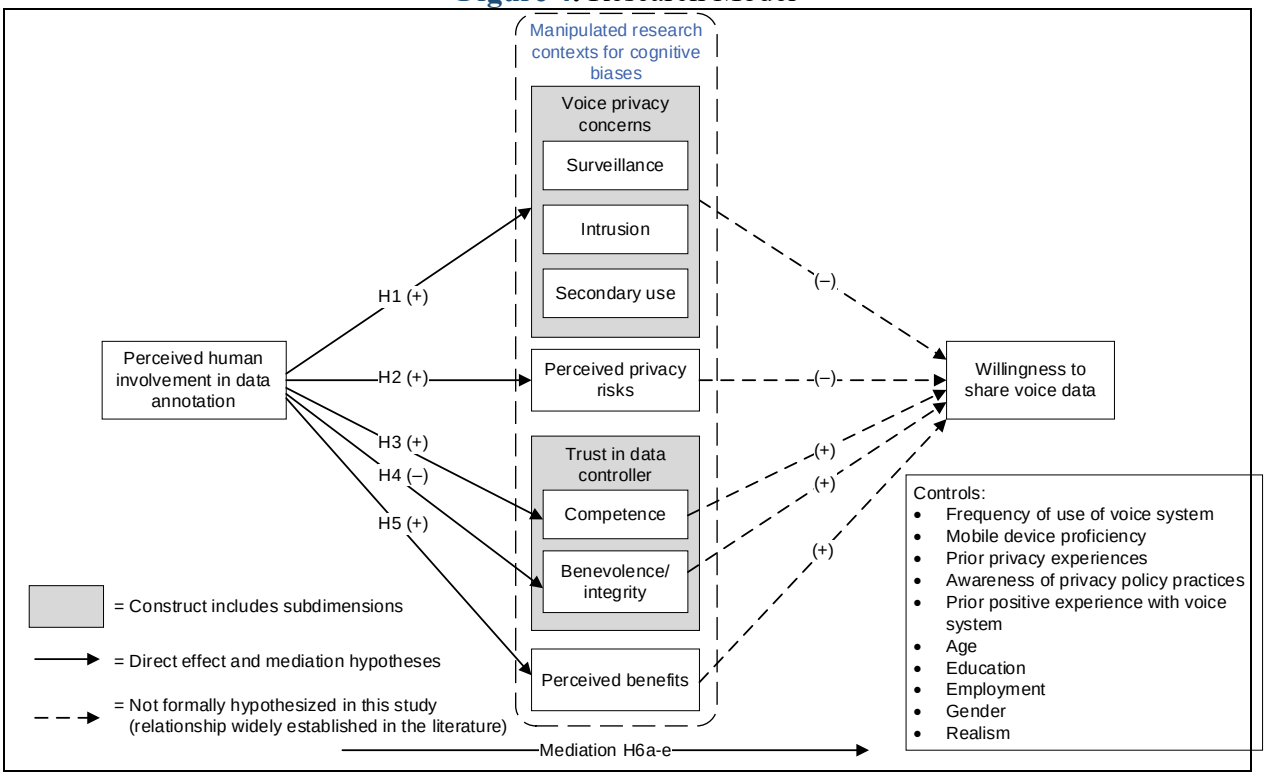
Moreover, researchers examined alternative explanations beyond the APCO, drawing from economic theories and incorporating principles from psychological experiments on decision-making and behavioral economics. This approach helped them better understand the effects of situational and contextual factors on consumers' privacy choices and actions (e.g., Acquisti, 2004, 2009; Acquisti & Grossklags, 2012; Adjerid et al., 2018a; Adjerid et al., 2018b; Tsai et al., 2011). The enhanced APCO theory was modeled to encompass explanations of people's privacy behaviors through two lenses: (1) the original APCO model (Smith et al., 2011) and the (2) boundary conditions related to "lower-effort cognitive processes and biases in human decision making" (Dinev et al., 2015, p. 640). We assume and build on these relationships and propose human involvement as an antecedent to the APCO constructs. Namely, we borrow the theory and explain how we recontextualize the APCO to investigate the types of cognitive biases and level of effort conditions that affect consumers' privacy decision-making when they view privacy policies.

1.2.4. Contextualizing the APCO for Human Involvement

The practice of borrowing and applying theories is prevalent in research (Whetten et al., 2009). Notably, horizontal borrowing is the study of phenomena in other types of social contexts. A primary reason to borrow theory is to apply the theory's explanatory power in a new context and to delineate its boundaries and scope conditions (Whetten et al., 2009). To borrow theory appropriately and to test primarily how the theory functions, the new and old contexts should be approximately equivalent (Morgeson & Hofmann, 1999). Furthermore, the context and level

sensitivities should be similar. Because the enhanced APCO is an individual-level theory and the context pertains to information privacy, namely, a person's behavioral or attitudinal reactions to a stimulus, the nomological relationships in this particular theory should hold true (Whetten et al., 2009). Our context involves consumers who interact with conversational agents as the stimulus, and we investigate their willingness to share training data as the attitudinal reaction. The focus of our study is to evaluate the role of human involvement in an ML context, and thus the use of a privacy theory that systematically explains all relationships between constructs is imperative to isolating the effects we intend to investigate, namely the operationalization of an antecedent to privacy concern and the contextual effects of cognitive biases. Hence, our aim is not to enhance the APCO theory itself, which would be regarded as a contribution *to* theory, but rather to apply it within the new context of human-machine learning augmentation, constituting a contribution *of* theory (Whetten et al., 2009). **Figure 4** depicts our proposed research model.

Figure 4. Research Model



1.2.5. Positioning Human Involvement as an Antecedent

We begin by establishing the context in which human involvement in ML systems is disclosed to consumers through privacy policies provided by companies. Companies seeking to comply with data privacy laws often use privacy policies to convey transparency in their data handling practices (Grant & Wischik, 2020). Privacy policies are based on "notice and consent" requirements and the assumption that consumers provide informed consent to permit companies to collect and process their data (Acquisti et al., 2013; Adjerid et al., 2018a). Research shows that consumers apply stereotypical thinking when they process privacy policies (Gerlach et al., 2019). Further, information asymmetries exist between companies and consumers, such that even if consumers fully comprehend the privacy policies they read, it would still be difficult to anticipate every downstream privacy risk involving their data (Acquisti et al., 2013; Acquisti et al., 2015; Adjerid et al., 2018b). Through the use of privacy policies, companies can thus attempt to appear open and transparent about the processing of consumers' data.

Awad and Krishnan (2006) explain that consumers make tradeoffs to improve the performance and personalization of the services they receive. Consumers provide their transparent, personal information for enhanced personalization, known as the "personalization privacy paradox" (Awad & Krishnan, 2006, p. 13). For example, Amazon Alexa's privacy policy frames *human involvement* as humans reviewing a *small* sample of requests to help Alexa understand the *correct* interpretations and to provide the *appropriate* responses in the future (Amazon, 2021; Merrill, 2020). The privacy policy refers to supervised machine learning as an *industry-standard practice*, and it explains that because Alexa works with a *diverse* range of customers that it is necessary for Alexa to respond properly to variations in customers' speech patterns, dialects, accents, and vocabulary. Human involvement is framed as a benefit to consumers and not as a risk, thereby influencing a consumer's privacy calculus toward the use of the voice system (Dinev & Hart,

2006). Formally, *privacy calculus* is the cost-benefit analysis people perform to determine whether they will accept a loss of privacy from any disclosure of personal information as long as they perceive the benefits of using the system to exceed their acceptable levels of risk (Culnan & Bies, 2003; Dinev & Hart, 2006). In the case of voice system use, if the benefits of using voice commands exceed the risks of disclosing personal voice data, then a user's privacy calculus will tilt toward a net positive outcome.

However, research shows that several cognitive biases and lower-levels of information processing (Alashoor et al., 2022) can cause consumers to misinterpret their reading of privacy policies or terms and conditions. First, most people skip over the privacy policy, or if they do read it, they do not spend enough time to fully extract the intended meaning of the content (Obar & Oeldorf-Hirsch, 2020). Second, consumers who do read privacy policies find them lengthy, inscrutable, and the language confusing (Milne & Culnan, 2004). As illustration, to read through every privacy policy a consumer encountered would take 25 days in a single year, and the estimated opportunity cost for them to read those privacy policies would be $781 billion per year (McDonald & Cranor, 2008). Thus, the sheer number of privacy policies a consumer would have to read can force users into taking mental shortcuts to apply general stereotypical thinking about the way companies' process consumers' data (Gerlach et al., 2019).

When consumers use voice agents, the cognitive biases and cognitive processing efforts are exacerbated by the reduced friction in the ease of use (Moriuchi, 2019) and the ability of designers to design the system to cultivate *trust* in an interpersonal relationship with the consumer (Nass & Moon, 2000; Nass et al., 1995; Rheu et al., 2021). Additionally, the absence of a visual interface in conversational ML systems for displaying privacy policies or terms and conditions makes it easier for users to bypass these documents entirely. This means that consumers must engage not

only with the conversational agent but also with the company's website to access its privacy policies (Howell, 2021).

We thus contextualize three types of cognitive effects when consumers interact with conversational ML systems. First, we argue that a salience bias is present. *Salience bias* is defined as favoring salient cues and stimuli over difficult, diffuse information, which can lead to suboptimal decision making (Lee et al., 2018). For example, salience bias can cause a person to be overconfident about the information presented explicitly before them, which is often insufficient for predicting an actual outcome (Griffin & Tversky, 1992). A mistaken assumption would be if a consumer interacts independently with a conversational agent when another human is *not* saliently present and thinks the interaction is confidential when it is actually not (Howell, 2021). We therefore operationalize salience bias through the covert or overt mentioning of human reviewers annotating consumers' voice data, as referenced in a company's privacy policy (Merrill, 2020). We define *covert* mention of human involvement as a privacy policy using generic terms such as "data analytics," "product development," "improvement," or "personalization" but not referencing the explicit review of personal data by human agents for algorithm training purposes (e.g., Facebook, 2021). *Overt* mention of human involvement indicates that the privacy policy specifically references human involvement in reviewing and labeling users' personal data for the training of algorithms (e.g., Amazon, 2021). In the absence of an explicit mention of human involvement, the mistaken assumption caused by salience bias would be that only machine processing of consumers' personal data will occur.

Second, default trust bias is inherent in privacy-decision making (Dinev et al., 2015). *Default trust bias* is defined as the innate instinctual heuristic people develop during infancy that lessens their conscious thought processing to evaluate the trustworthiness of others (Dinev et al., 2015).

We thus posit that when consumers believe that ML systems operate similarly to traditional IS through stereotypical thinking (Gerlach et al., 2019), they adopt a "tool" view toward a conversational agent and believe it does *not* have its own agency. Because the consumer feels in control of the interaction, the control paradox explains that consumers may generate a mistaken belief of privacy protection toward their interactions with the system (Brandimarte et al., 2012). Furthermore, if consumers have a satisfactory experience with a product or service from a company they trust, the default trust will transcend across the company's full product offerings through brand trust (Chaudhuri & Holbrook, 2001). Therefore, a consumer may, by default, trust that companies are complying with data privacy laws and are respecting consumers' information privacy (Dinev et al., 2015; Gerlach et al., 2019).

Finally, *low level of effort in cognitive processing* is defined as "relatively little cognitive effort or conscious awareness" (Dinev et al., 2015, p. 643). The rational model of privacy calculus assumes that privacy-related behaviors occur from deliberate, high-effort cognitive processing; however, research explains that consumers apply cognitive heuristics and mental shortcuts to avoid "effortful analysis involving logic and elaborated reasoning" (Dinev et al., 2015, p. 643). Namely, consumers may not consider the downstream consequences of human involvement and the training of their voice data and instead rely on cognitive shortcuts to focus on the salient benefits of using conversational agents.

1.2.6. Human Involvement and the "Privacy" Tradeoff

Privacy concerns arise when a person feels incapable of taking action to protect oneself. (Li, 2012). Predominantly, people are concerned with seven types of privacy: privacy of the (1) *person*, (2) *behavior and action*, (3) *communication*, (4) *data and image*, (5) *thoughts and feelings*, (6) *location and space*, and (7) *association (multilevel privacy)* (Finn et al., 2013). We propose that conversational ML systems are capable of infringing on the seven types of privacy, especially

through human involvement and the contextualization of consumers' voice data. First, conversational agents collect and store verbal communication (communication privacy). Second, developers build smart products with cameras and other sensors to collect personal data about people inside their homes (data and image, location and space, and person privacy) (Menard & Bott, 2020). Third, through human involvement, human agents can infer a person's thoughts, feelings, behaviors, and actions, and can identify those who are in company with the data subject (behavior and action, thoughts and feelings, and association privacy) (Ågerfalk, 2020; Zuboff, 2015, 2019). Because consumers may be unaware of human involvement in the reviewing of their data, they may not initially feel helpless in protecting themselves. However, when human involvement is overtly disclosed, we argue that consumers will feel incapable of protecting their information and experience heightened levels of privacy risk. Moreover, because they have already consented to terms and conditions, they will inevitably not know with whom their data are shared (Howell, 2021). Thus, we posit that:

H1. *Perceived human involvement* is positively associated with *privacy concern*.

Conversational agents actively await wake words and are thus persistently recording (Cox, 2019; Day et al., 2019; Howell, 2021). ML systems are characterized as being *aware* of their environments with the data subjects *unaware* of their use of the device (Howell, 2021; Kane et al., 2021; Schuetz & Venkatesh, 2020). When consumers are unaware of the ways companies can collect and use their data, an information asymmetry arises (Acquisti et al., 2020). This, in turn, increases a company's ability to collect more consumer information. Because of the information asymmetry, consumers cannot respond to any risks to which they are unaware (Acquisti et al., 2020). However, the overt explanation of human involvement allows consumers to envision the ways their data may be used, such as unnecessary reidentification and aggregation risks (Zuboff, 2019), and can activate a new calculus toward risk and invite consumers to consider new risk

mitigation strategies due to changes in risk perceptions (Adjerid et al., 2018a). To this end, we posit that:

H2. *Perceived human involvement* is positively associated with *privacy risk*.

1.2.7. Human Involvement and the "Performance" Tradeoff

Trust is generally accumulated during the social relationship between companies and consumers (McKnight et al., 1998). Consumers who have a high opinion toward technology reflect "confidence or optimism regarding adoption of new ideas or technologies" (McKnight et al., 2002, p. 340). ML systems are novel and their capabilities are continually expanding (Kane et al., 2021; Teodorescu et al., 2021). Research shows that people with a proclivity toward novel technologies have a higher disposition to trust the product, service, or company (McKnight et al., 2002). Also, when companies provide detailed privacy policies to explain their products and services, people have a higher tendency to trust the company (Wu et al., 2012). Companies can thus use privacy policies and terms and conditions to exhibit an image of fairness and transparency in their ML systems (Grant & Wischik, 2020) and elicit higher levels of trust. Clearly explaining their data collection and handling practices, companies can thus engender trust in the overall system and demonstrate that their conversational agents are unbiased, continually trained, inspected, and improved. When companies describe their intentions to improve a product with the use of consumers' personal data to add personalization services (Sutanto et al., 2013), companies can further highlight the competency of their systems.

H3. Perceived human involvement is positively associated with competence trust.

Nevertheless, if a company openly reveals human involvement but fails to clarify the potential risks associated with human agents accessing users' personal information, the privacy policy may cease to serve as an effective instrument for transparent disclosure. Instead, it serves as an illusory form of transparency, causing the consumer to solicit more information to understand how their

data was and will be used (Grant & Wischik, 2020). Because organizational best practices advocate for human augmentation in ML practices (Grønsund & Aanestad, 2020; Murray et al., 2020), we believe consumers will similarly see the improved performance of conversational agents when human involvement occurs. Also, companies frame human involvement as a benefit to consumers through the improved performance and personalized services of their conversational agent. However, if a privacy policy lacks transparent disclosure regarding the direct involvement of human agents, and consumers discover this information through external sources, it can lead them to question the company's commitment to their best interests and perceive a lack of honesty. Thus,

H4. Perceived human involvement is negatively associated with benevolence and integrity trust.

Privacy calculus is the assessment consumers perform when they decide to disclose their information that it will be used fairly and they will not be exposed to negative consequences as a result of the disclosure (Awad & Krishnan, 2006; Culnan & Armstrong, 1999; Dinev & Hart, 2006). The difference of this calculus is considered perceived value (Zeithaml, 1988). A positive perceived value is considered the benefit of information disclosure (Dinev & Hart, 2006). In the context of conversational agents, we define perceived value as the overall utility a consumer receives when furnishing their voice data to improve the conversational interaction with the system. Because conversational agents operate through a voice interface, the consumer may not understand how their data are being processed and shared (Felt et al., 2012; Menard & Bott, 2020). Thus, the potential privacy invasion is not concrete (Menard & Bott, 2020). Moreover, the presence of a perceived risk may subtly manifest itself through the occasional utterances by the conversational agent when it suddenly awakens without the consumer's prompting (Merrill, 2020). However, the longer consumers experience the benefits of use, this may amplify their perceptions of benefits over perceived risks (Xu et al., 2011b). Thus, when companies frame human

involvement as a benefit to consumers through the improved performance of the conversational agent, consumers will perceive an increased value in human involvement. Hence,

H5. Perceived human involvement is positively associated with perceived benefits.

The APCO proposes that privacy concern, privacy risk, trust, and benefits are deliberate considerations people undertake before performing information disclosure or sharing behaviors (Dinev et al., 2015; Smith et al., 2011). Antecedents to privacy concern are thus fully mediated through the APCO constructs. We follow the APCO in emphasizing the mediating roles of the APCO constructs and the inherent deliberate considerations that flow from antecedents to privacy concern to outcomes. Therefore,

H6. (a) Privacy concern, (b) privacy risk, (c) competence trust, (d) benevolence and integrity trust, and (e) perceived benefits will mediate the effect of **perceived human involvement** on **willingness to share voice data**.

1.3. Research Method

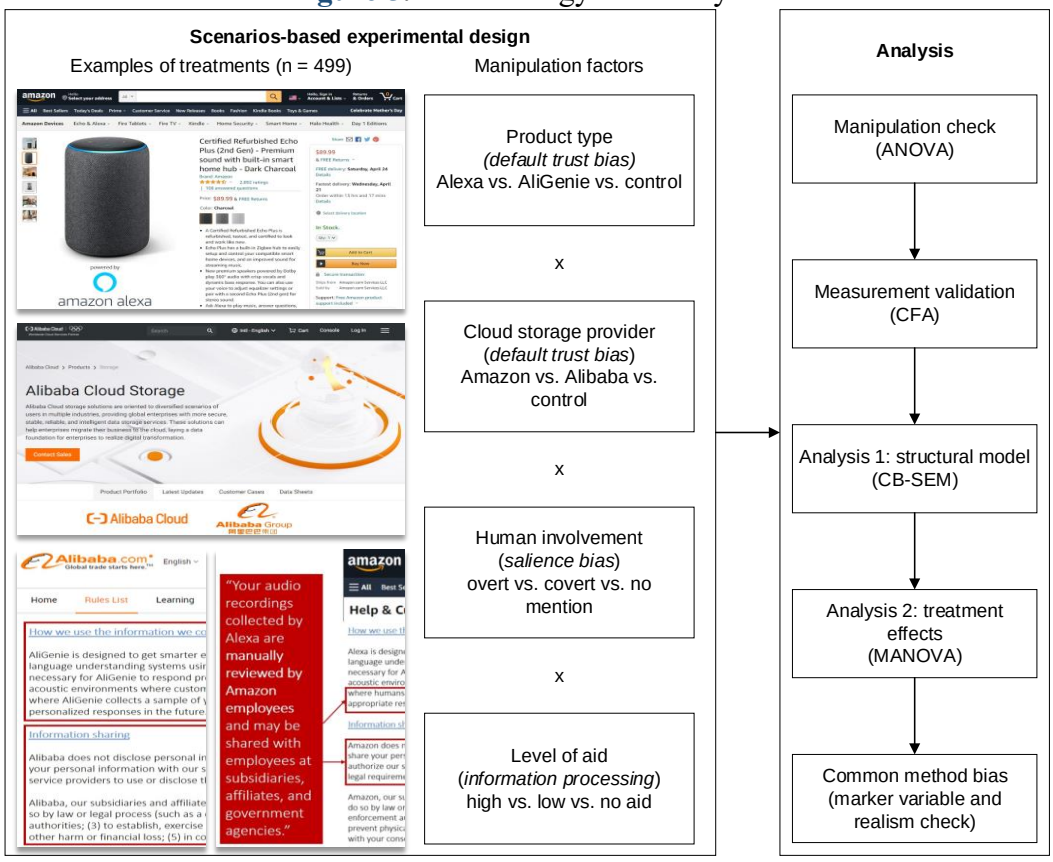
To test our hypotheses, we used a scenario-based survey with a 2 x 2 x 2 x 2 experimental orientation to present vignettes of graphical representations—containing realistic details of privacy policy situations—to collect participants' subsequent responses to measurement items on our dependent variable of interest (Trevino, 1992). Scenario-based surveys are commonly used to investigate decision-making in hypothetical situations, and the vignettes serving as treatments often contain concrete information on the independent variables (Trevino, 1992). Notably, using vignettes allows for the creation of controlled, situational contexts that can be experimentally varied during participants' decision-making tasks. Thus, we can create multiple contextually rich vignettes and embed them into an experimental design to enhance the realism of the decision-making situation (Vance et al., 2015) and to manipulate factors of theoretical interest (Al-Natour et al., 2020). By varying the contextual details scripted and presented in each of the descriptive vignettes, we can build a factorial design that includes a combination of specific levels of factors

of interest to create multiple treatment groups from which to hypothesize and investigate effects of several factors on the judgments, preferences, and decisions of survey participants (Rungtusanatham et al., 2011). **Figure 5** depicts examples of the privacy policy treatments we administered in our scenario-based experimental design, the four factors we manipulated on each of the privacy policies, and the corresponding analyses we performed on the collected data.

1.3.1. Scenario Designs and Manipulations

To enhance ecological validity, we closely emulated the content furnished in privacy policies of leading technology companies and provided screen captures of actual company webpages, products, and services to ensure our vignettes depicted realistic content that participants would encounter during their typical online e-commerce experience. Amazon is rated as the most trusted company in America (Seitz, 2021), and we used them as our company of interest.

Figure 5. Methodology and Analysis Process



Moreover, to mitigate effects of default trust bias, we also used Alibaba because of its leadership in providing voice products and services in China (Sun et al., 2021a). We developed realistic vignettes depicting actual Amazon and Alibaba products and services and carefully adapted the wording in excerpts from privacy policies from Ring, Samsung, and Amazon, and embedded these into vignette sets in a scenarios-based survey. Finally, as an added measure, we included a realism item as a control for the effects of scenario realism (Siponen & Vance, 2010). Realism items are important to ensure causal realism in the relationship between treatments and sets of variables (Straub et al., 2004).

We created three primary sets of vignettes to cultivate distinct user experience conditions for three types of consumers: ones who are asked to envision interacting with (1) an Amazon Echo (Alexa), (2) an Alibaba Tmall Genie (AliGenie), or (3) a generic smart speaker and conversational agent without any affiliation to a specific product or company name (control). The Amazon Echo (Alexa) vignettes included a product page of an Amazon Echo speaker powered by Alexa, along with three pages of privacy policies that included the Amazon header across each page. Similarly, the Alibaba Tmall Genie (AliGenie) vignettes included a product page of an Alibaba Tmall Genie powered by AliGenie, along with three pages of privacy policies with Alibaba's header emblazoned across each. The third set of vignettes is the control and neither included a picture of a smart speaker nor any company header across each of the three privacy policy pages; instead, the smart speaker and conversational agent were described in written text without any specific reference to a product or cloud service provider type. The purpose of the control condition is to detect the treatment effect based on the type of product participants are asked to review.

Moreover, as our focus is on users' views about companies' data handling practices in general, rather than their usage of a specific smart device brand, we incorporated an additional factor to

examine default trust bias: the cloud storage provider responsible for storing voice data collected by the conversational agent. Formally, conversational agents retrieve responses to user commands from cloud servers (Amazon, 2021), and the largest cloud providers rendering requests from Amazon and Alibaba conversational agents are Amazon Web Services and Alibaba Cloud, respectively. We thus included an additional manipulation in the three sets of vignettes, one for Alibaba Cloud and Amazon Web Services cloud storage provider types. In summary, each set of vignettes consisted of pertinent information that could potentially influence a participant's willingness to share training data with a company. The sets namely included (1) a product page of a smart speaker, which were actual screenshots from Amazon.com and Alibaba.com (translated into English); (2) a cloud service provider page, which included actual screenshots of services from AWS.com and Alibabacloud.com; and (3) three pages of privacy policies illustrating actual content obtained from Samsung's, Ring's, and Amazon's privacy policies that included small, yet carefully worded modifications to the texts.

The careful alterations to the privacy policy included our (1) replacing company names of original policies with the companies of interest in our survey, namely Amazon and Alibaba; (2) altering the "covert" description of human involvement in Amazon's description of Amazon Echo and Alexa services (because Amazon's original statement on human reviewers is used as the "overt" description) (Amazon, 2021); and (3) deleting content from the original privacy policies to more prominently highlight the written policy on the use of consumers' training data, specifically that consumers' voice data are collected, reviewed, and shared by companies as explicitly referenced in privacy policies (i.e., "Information we collect," "How we use your information," "Information sharing" sections). Appendix 1C shows the images of the treatments and the manipulated factors included in the experiment.

Because we are primarily interested in privacy concerns related to perceived human involvement in ML systems, we made extensive efforts to make human involvement salient in the minds of participants by including two factors specifically related to it. First, to overcome salience bias, we used clear language in the privacy policy to convey the activity of human agents in the processing of users' voice data, a term we phrase as the covert vs. overt mention of human reviewers. Second, we provided aids to participants to facilitate their information processing of the privacy policy content, because research demonstrates information asymmetries exist and consumers may not systematically process the content in privacy policies (Acquisti, 2009; Acquisti et al., 2015; Dinev et al., 2008; Gerlach et al., 2019). Thus, to improve the level of effort in information processing, we provided two levels of aids: a high-level condition where participants are given clear explanations on the meaning of the privacy policy content, and a low-level condition where participants receive only callout boxes to serve as visual cues to heighten awareness of important privacy policy content. Details on each of the four factors vary among the products, services, and privacy policy pages. Consequently, we created a scenarios-based experiment with a 2 (*salience bias*: human involvement) x 2 (*default trust bias*: product type) x 2 (*default trust bias*: cloud service provider) x 2 (*level of effort*: aid rendered) between-subjects factorial design to form 16 treatment conditions, and an additional group as a control condition, for a total of 17 treatment groups.

1.3.2. Procedures

We used the Qualtrics platform to create our survey and advertised the study as a website readability and product usability study on users' evaluation of smart speaker devices. This was to prevent possibly priming participants of our study relating to a privacy context (Al-Natour et al., 2020). Specifically, we told participants to imagine they recently purchased a smart speaker device (i.e., Amazon Echo, Alibaba Tmall Genie, or control) and to evaluate a series of lengthy, complex

privacy policies before assessing their willingness to use the device. The instructions informed participants to evaluate a series of five screenshots taken from the websites of large online commerce companies that sell smart speaker devices that recognize their voices. Their task was to carefully consider the designs of the websites and the smart speakers the companies were selling. They were then told we would ask for their impressions and opinions about their product.

Because our study relates to voice data collected by conversational agents like Siri, AliGenie, Cortona, and Alexa, we screened participants based on their familiarity with and use of conversational agents. We provided a presurvey for screening participants and collected items relating to the marker variable and their disposition to trust. If participants were unfamiliar with any kind of virtual assistant or never used a smart speaker device, they were unable to participate in our usability study. After initial screening, remaining participants were randomly assigned to one of 17 conditions and were presented with a set of Amazon, Alibaba, or control vignettes, followed by a post experiment questionnaire. After viewing the product and cloud provider pages and carefully reading the privacy policy excerpts, participants answered four attention questions related to the experimental cues information contained in the vignettes. Their responses were validated against the product type, cloud provider, human involvement, and level of aid treatments embedded in their specific treatment condition. Correct responses allowed participants to proceed to the post experiment. However, if participants answered incorrectly to any one of the four attention checks, they were presented with the treatments a second time and allowed to answer the attention questions. If they again responded incorrectly, they were screened from the study.

In the post experiment, we asked participants their willingness to share their voice data to train speech recognition software and to improve the company's products and services. We also asked an open-ended question for them to share their specific reasons why they would or would not share

their voice data with the company. Next, we asked participants to respond to measurement items on perceived human involvement to serve as a psychometric manipulation check on human involvement to verify whether we successfully manipulated factors related to participants' observing a company's transparent explanation on the types of agents (voice agent vs. human agent) that process consumers' data. Because product type, cloud service provider, and level of aid were saliently provided, with the respective company's header emblazoned on all screenshot pages or a vivid explanation of the privacy policy in a red-colored box, we did not ask any psychometric measurement items on those factors. Instead, we allowed the attention checks of selecting the correct company's product and service and amount of aid we provided to help with their information processing to suffice as clear and evident manipulation checks. Next, we included items that captured participants' trust toward data controllers, privacy concerns, privacy risks, and perceived benefits. Finally, we collected demographic information, items relating to their experience with smart devices, and control variables related to privacy policy awareness and prior privacy experiences. The APCO model includes awareness and privacy experiences as antecedents to privacy concern; therefore, we included these variables as controls (Dinev et al., 2015). After participants completed the questionnaire, we debriefed them about our intention of the study to investigate training data privacy concerns and asked for their formal consent to use their responses in our analysis. The study was formally approved by the Institutional Review Board.

1.3.3. Construct Measures, Item Development, and Pilot Testing

We measured all constructs using multi-item 7-point Likert-type scales using an assortment of scale endpoints ranging from "strongly disagree to strongly agree," "describes me extremely poorly to describes me extremely well," "extremely unbelievable to extremely believable," "never to always," and "does not describe my feelings to completely describes my feelings" (Appendix 1B includes the survey instrument and lists the measurement details). Importantly, positivist

studies primarily use 7-point Likert-type scales (Joshi et al., 2015), and the use of multi-item scales is most common to validly and reliably measure constructs in a questionnaire (Robinson, 2018).

We developed measurement items based on an extensive review of the privacy and trust research discourses, and adapted items from scales that had been validated in earlier studies. We situated each item in the context of conversational agent use to capture participants' perceptions more accurately on privacy and trust during their interactions with conversational agents. The demographics and smart device use questions were obtained from a variety of sources (Al-Natour et al., 2020; Crossler & Bélanger, 2019; Lopatovska et al., 2018; Malhotra et al., 2004) and administered in the post survey in either original or slightly adapted forms. The scales measuring disposition to trust in the presurvey and trusting beliefs in the post survey were adapted from Moody et al. (2014) and Moody et al. (2017), with the items specifically adapted to reflect trust toward the company handling users' voice data and not toward "the seller," as was referenced in the original context of the items. Response set items such as "If you are not a fish, then select "Agree" as the response to this question" were included in the pre and post surveys. Last, the blue attitude marker scale from Miller and Chiodo (2008) was retained in original form and administered in the presurvey.

In the post-experiment survey, we adapted the willingness to disclose information scale from Moody et al. (2017) as the dependent variable. Notably, because the contexts of information sharing in the original scale and our adapted scale are vastly different, the adapted items are completely contextualized to reflect participants' willingness to share voice data: to improve voice recognition services, for data analysis purposes, and to receive personalized services. In contrast to the original items, which pertained to sharing specific types of personal information such as name, address, social security numbers, and product needs. Privacy concerns was adapted from

Xu et al.'s (2012a) mobile users' information privacy concerns scale to reflect a shift in focus from mobile apps to Alexa and AliGenie. Finally, items on privacy risks were adapted from Hong and Thong (2013), including a shift in focus from websites to Alexa and AliGenie, and items on perceived benefits were adapted from Xu et al. (2011b), written to reflect Alexa and AliGenie rather than the M-Coupon service. Additional scales measured in the post survey that served as control variables included items for: prior privacy experience, adapted from Smith et al. (1996) to reflect data collection from Alexa and AliGenie rather than "the Internet"; awareness of privacy practices, retained from Malhotra et al. (2004); and past positive experience, adapted from Pavlou and Gefen (2004) with a change in reference from Amazon's auction marketplace to Amazon Alexa and Alibaba AliGenie.

In addition, we developed a psychometric measure to evaluate our human involvement manipulation. Because we could not find existing scales in the extant human-ML systems augmentation research that measured the level of human involvement in human-in-the-loop or human-centered AI configurations, we adapted original items from a variety of studies (Boh & Yellin, 2006; Lassen et al., 2007; Sasidharan et al., 2011; Segaar et al., 2007) in two specific contexts: (1) enterprise systems and architectures where the objective was to measure stakeholder involvement in the implementation of enterprise systems across an organization, and (2) health campaign program development where researchers investigated employee involvement in organizational health campaigns. Items from both contexts were adapted to reflect the extent to which participants believed human reviewers were involved in the training of the speech technologies used in conversational agents. The human involvement measures were then carefully inspected by the authors to ensure content and face validity.

After carefully adapting the measurement scales in our survey, we conducted a pilot test with

74 complete and valid responses from Amazon's Mechanical Turk (MTurk) platform. Importantly, we tested the perceived human involvement items to analyze the content and discriminant validities and the results confirmed the scale as both valid and reliable. Moreover, our statistical analyses of the pilot data showed that the instrument performed as expected and required only minor corrections, which allowed us to proceed with the full data collection.

1.3.4. Final Data Collection

The experiment included 499 responses and was conducted online using participants recruited through Cloud Research from the MTurk platform. Research shows that responses collected voluntarily from MTurk can be reliable sources of data when reasonable data-quality controls are used (Buhrmester et al., 2011; Lowry et al., 2016; Mason & Suri, 2012; Steelman et al., 2014). We followed recommended recruiting and screening techniques for crowdsourcing platforms. To ensure high quality data, we specified in the Cloud Research selection criteria to solicit responses from MTurk workers with greater than 97% HIT approval rates and number of approved HITs exceeding 100. Additionally, we restricted access to our survey to workers in the United States. Finally, we applied procedural controls that included attention and comprehension checks, explaining the importance of the study in the instructions, providing a warning that inattentive respondents would not receive compensation, and obtaining a large sample for the final study (Hulland & Miller, 2018; Lowry et al., 2016; Steelman et al., 2014). The compensation for a completed survey was $2.00, and the expected time to complete was approximately 20 minutes.

Among the participants, 47.9% were female, with an age distribution for all participants between 18 to older than 65 years old. On average, the participants took 20.68 minutes to complete the survey. All participants were familiar with or used virtual assistants with 72% of participants using a voice assistant at least once a day. Most participants (49.9%) owned a smartphone with the iOS operating system, followed by 49.1% of participants who owned an Android device. Almost

all participants had never used AliGenie (Mean = 1.86, SD = 1.52), and those who had used Alexa rated their quality of experience between average and good (Mean = 4.562, SD = 1.12). Nearly all participants strongly agreed that companies should conspicuously disclose their data handling procedures (Mean = 6.31, SD = 0.81), and participants seldom experienced prior privacy invasions (Mean = 3.01, SD = 1.24). A complete summary of the descriptive statistics of the participants is shown in Appendix 1D.

1.4. Analysis and Results

We performed our analysis in several steps (see **Figure 5** above). First, we ascertained the validity of our research design. Second, we evaluated whether our randomization was successful by ensuring our experiment had a balanced design with approximately similar numbers of participants in our treatment cells. Next, we performed a manipulation check to ensure the desired effects were achieved, then we examined the reliability and validities of our survey instrument. To test the effect of perceived human involvement on the APCO constructs, we conducted a covariance-based structural equation modeling (CB-SEM) analysis (H1–H5). Then, to test whether the APCO constructs mediated the effect of perceived human involvement on willingness to share, we performed bootstrapped confidence interval (CI) tests for full and partial mediation (H6a–e). To test the contextual effects of our operationalized cognitive biases, namely salience bias (H1–H6), default trust bias, and information processing, we conducted a multivariate analysis of variance (MANOVA). Finally, we performed various methods to detect whether the presence of common method bias (CMB) influenced our findings. In the following sections, we detail the procedures undertaken in each of these steps.

1.4.1. Randomization, Assignment to Experimental Groups, and Manipulation Check

In our research design, we employed a 2 (human involvement) x 2 (product type) x 2 (cloud provider type) x 2 (level of aid) factorial design. Participants were randomly assigned to one of 17

experimental groups using the Qualtrics platform (**Table 1**).

Table 1. Participant Distribution by Randomized Grouping (n = 499)

		Research Context Manipulations			
Condition number	Human involvement	Voice system	Cloud provider	Level of aid provided	n per cell
1	Covert	Alexa	Amazon web services	High	29
2	Covert	Alexa	Alibaba cloud	High	31
3	Covert	Alexa	Amazon web services	Low	29
4	Covert	Alexa	Alibaba cloud	Low	23
5	Covert	AliGenie	Amazon web services	High	28
6	Covert	AliGenie	Alibaba cloud	High	26
7	Covert	AliGenie	Amazon web services	Low	35
8	Covert	AliGenie	Alibaba cloud	Low	26
9	Overt	Alexa	Amazon web services	High	31
10	Overt	Alexa	Alibaba cloud	High	27
11	Overt	Alexa	Amazon web services	Low	26
12	Overt	Alexa	Alibaba cloud	Low	23
13	Overt	AliGenie	Amazon web services	High	33
14	Overt	AliGenie	Alibaba cloud	High	30
15	Overt	AliGenie	Amazon web services	Low	26
16	Overt	AliGenie	Alibaba cloud	Low	27
17	No mention	No mention	No mention	None	49

Next, manipulation checks are needed to provide additional evidence that the manipulated factor (X) has an influence on the dependent variable (Y) (Bagozzi, 1977). Moreover, manipulation checks verify that the actual manipulation corresponds with the manipulation effect the researchers intended to observe on the basis of the theory guiding their research design (Boudreau et al., 2001). A manipulation check was needed for the psychometric measure of human involvement perceptions. We verified that participants had received the manipulation and that the privacy policy descriptions led to differences in perceived human involvement in supervised machine learning practices.

Participants were required to answer four comprehension checks, related to each of the four manipulations, to ensure they recognized the condition to which they were assigned. Namely, for those in the "overt" human involvement condition, participants were required to acknowledge that human reviewers would review users' voice data by correctly selecting "human reviewers" in the

comprehension check. Conversely, participants in the "covert" condition and "control" conditions were not told explicitly that humans were involved and thus participants would either select that the device trained itself or that there was no mention of machine or human reviewers in the ML training process, respectively. To test for construct validity, that is, whether the human involvement manipulation corresponded with the intended manipulation, we adapted "involvement" items from existing research discourses to develop manipulation check questions. **Table 2** shows that our human involvement manipulation led to significant differences in participants' responses among the three groups and was therefore successful.

Table 2. Manipulation Check for Human Involvement (n = 499; df = 496)

Manipulation	Mean (SD)			F-statistic	p-value	Manipulation successful?
	Overt	Covert	Control			
Human involvement	5.40 (0.95) (n = 223)	4.21 (1.38) (n = 227)	4.15 (1.17) (n = 49)	63.765	$p < .001$	Yes

We used confirmatory factor analysis to evaluate the measurement model and to determine factor loadings (Straub et al., 2004). The factor analysis showed that all item loadings exceeded the 0.60 minimum (Hair et al., 2006). To assess convergent and discriminant validities, we computed correlations between latent constructs, descriptive statistics, reliabilities, average variance extracted (AVE), and the square root of the AVE. **Table 3** and **Table 4** summarize the construct-level statistics.

For each construct, the Cronbach's alpha was above the 0.70 minimum, and the average variance extracted (AVE) was above 0.50 (Gefen et al., 2000), thus indicating convergent validity. Discriminant validity was also supported, such that the square root of AVE for each construct was greater than its correlation values (Anderson & Gerbing, 1988). The composite reliability for each construct was above the recommended minimum of 0.70 (Fornell & Larcker, 1981). We evaluated multicollinearity by estimating the variance inflation factors (VIFs), and all VIFs were below the

recommended 4.0 threshold (Aiken & West, 1991). Based on these findings, the measurement model meets established standards.

Table 3. Descriptive Statistics, Reliabilities, and Multicollinearity

Construct	Mean	SD	CR	CA	AVE	VIF
Awareness	6.31	0.81	0.828	0.817	0.620	1.256
Benefits	5.23	1.16	0.877	0.861	0.706	1.843
Benevolence/integrity	4.05	1.36	0.950	0.949	0.733	2.196
Competence	5.17	1.17	0.911	0.910	0.719	1.999
Concern	5.17	1.48	0.970	0.969	0.781	3.682
Human involvement	4.74	1.33	0.923	0.926	0.705	1.126
Marker	5.24	1.09	0.820	0.792	0.605	1.063
Privacy experience	3.01	1.24	0.837	0.825	0.636	1.295
Realism	5.63	1.36	n/a	n/a	n/a	1.112
Risk	4.61	1.55	0.950	0.950	0.825	3.314
Share	3.84	1.81	0.920	0.919	0.793	2.001

Note. AVE = Average variance extracted; CA = Cronbach's alpha; CR = Composite reliability; SD = Standard deviation; VIF = Variance inflation factor; n/a = not applicable. Realism is a single item that was asked at the conclusion of the experiment phase to assess participants' perceptions of the realism of the privacy policy treatment.

We used several statistical and procedural remedies to address the potential concern of CMB (Podsakoff et al., 2003). For procedural remedies, we included instructions to participants to answer honestly and transparently because their answers were critical to our study. We informed them they would not be compensated if they failed to respond to the survey carefully. Furthermore, we reassured them of their anonymity. Next, we informed them that our research involved the evaluation of actual product web pages and privacy policies of leading technology companies, then asked them a realism item at the end of the experiment phase to assess their perception of the treatments. In the survey, we randomized the sequence of the questions and provided participants with different response formats to measure the dependent variable: (1) quantitative Likert-type scale format and (2) open-ended comment to explain their reason for their selection.

During the data analysis stage, we performed robustness analyses to detect the presence of CMB in our results. First, we conducted an exploratory factor analysis and applied Harman's single-factor test (Harman, 1967). Namely, we used a principal component analysis for all the items in the 14 first-order latent variables we measured in our study and found 14 factors with

eigenvalues greater than 1 that explained a cumulative variance of 84.94% of the total variance. The first factor accounted for only 34.48% of the total variance, which indicates a lack of a substantial CMB. Second, we used the latent marker variable (MLMV) approach to examine and partial out potential CMB (Lindell & Whitney, 2001; Malhotra et al., 2006; Podsakoff et al., 2003). Specifically, we included a theoretically unrelated "blue marker" variable (Miller & Chiodo, 2008) and measured it using a three-item scale. We followed the method of MLMV with a construct-level correction because research demonstrates that MLMV can reduce the confounding effect of CMB on structural path estimates by about 72% (Chin et al., 2013). Moreover, the effects of the marker variable were statistically nonsignificant when we included the items into our model as a predictor to willingness to share. We are thus confident that our study is relatively robust against CMBs.

Table 4. Correlations between Latent Constructs

Construct	1	2	3	4	5	6	7	8	9	10
1. Awareness	.788									
2. Benefits	.158	.840								
3. Benevolence/integrity	-.112	.518	.856							
4. Competence	.197	.630	. 597	.848						
5. Concern	.295	-.313	-.553	-.326	.884					
6. Human involvement	.186	.064	-.032	.120	.253	.840				
7. Marker	.099	.199	.189	.158	-.019	.058	.778			
8. Privacy experience	-.064	-.225	-.313	-.290	.436	.112	-.046	.798		
9. Risk	.213	-.367	-.562	-.444	.845	.189	-.099	.408	. 908	
10. Share	-.190	.436	.583	.370	-.643	-.120	.092	-.210	-.634	.890

Note. Diagonal elements in bold are the square root of AVE.

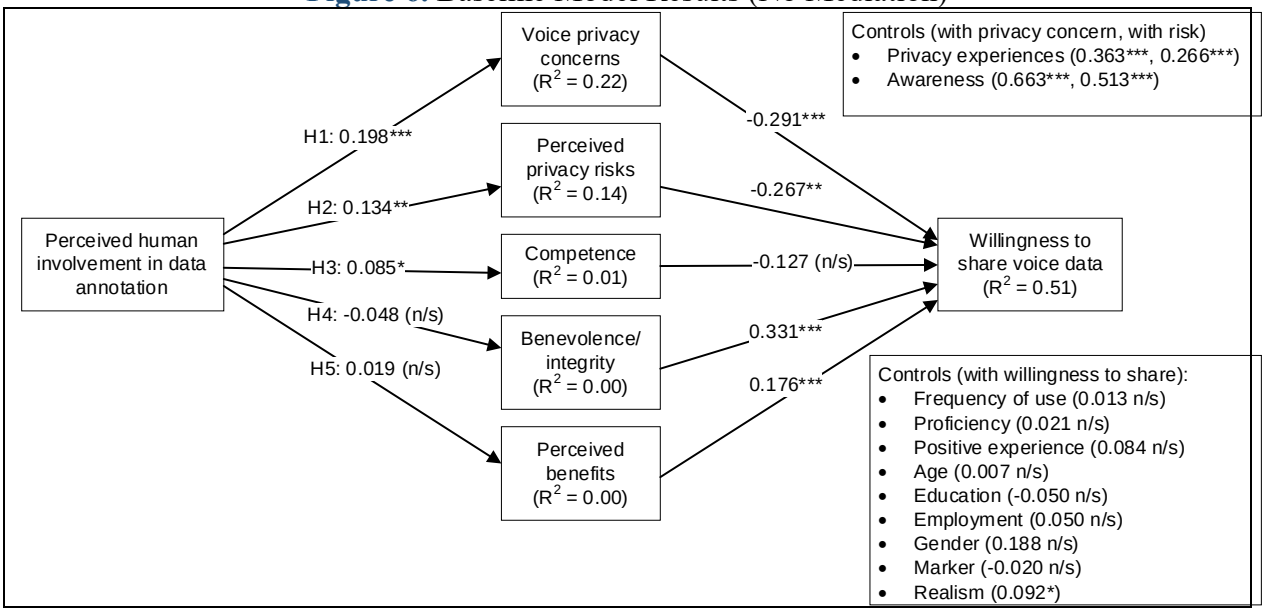
1.4.2. Main Effects Model and Mediation Testing

We tested the model with AMOS 26 and used covariance-based structural equation modeling. We tested H1 through H5 using a baseline model without mediation testing. **Figure 6** and **Table 5** depict the results of our hypothesis testing and show that most hypotheses are supported in our structural model. We added covariates to the model to test potential counter explanations of the results. Prior privacy experiences and awareness of privacy practices are validated antecedents to

privacy concern in the APCO model. Both covariates were significantly related to privacy concern and privacy risk. Next, the perceived realism of the privacy policies had a significant effect on how participants rated their willingness to share their voice data. The remaining covariates (i.e., frequency of use, mobile proficiency, past positive experience, age, education, employment, gender, and marker variable) were not significant.

Next, we assessed the model fit of our estimated model. The standardized root mean square residual (SRMR) score of .0764 is below the restrictive cutoff at < 0.080, which represents a good fit to the data (Henseler et al., 2016). The root mean squared error of approximation (RMSEA) of 0.043 was below the recommended threshold of 0.06 (Hu & Bentler, 1999). The comparative fit index (CFI) and the Tucker-Lewis index (TLI) were above their recommended minimums of 0.95 (Rigdon, 1996) and 0.95 (Tucker & Lewis, 1973), respectively . In summary, our measurement model demonstrated good model fit, good convergent validity, good discriminant validity, good reliability, and a lack of multicollinearity.

Figure 6. Baseline Model Results (No Mediation)



Note. *** = $p < 0.001$, ** = $p < 0.01$, * = $p < 0.05$, n/s = not significant; the model fit meets recommended heuristics: CMIN / *df* = 1.907, CFI = 0.961, NFI = 0.921, IFI = 0.912, TLI = 0.956, RMSEA = 0.043, PCLOSE = 1.000, SRMR = .0764; mediation effects predicted in H6a–e cannot be tested with the baseline model.

Table 5. Detailed Results of Baseline Model Testing

Relationship	β	SE	*t*-statistic	*p*-value	Hypothesis supported?
H1. Human involvement → Concern	0.198	0.049	4.009	< 0.001	H1 supported
H2. Human involvement → Risk	0.134	0.047	2.831	0.005	H2 supported
H3. Human involvement → Competence	0.085	0.036	2.359	0.018	H3 supported
H4. Human involvement → Benevolence/integrity	-0.048	0.048	-1.008	0.313	H4 not supported
H5. Human involvement → Benefits	0.019	0.049	0.396	0.692	H5 not supported
Concern → Willingness to share	-0.291	0.078	-3.727	< 0.001	n/a
Risk → Willingness to share	-0.267	0.089	-3.005	0.003	n/a
Competence → Willingness to share	-0.127	0.082	-1.539	0.124	n/a
Benevolence/Integrity → Willingness to share	0.331	0.061	5.46	< 0.001	n/a
Benefits → Willingness to share	0.176	0.037	4.779	< 0.001	n/a
Covariates					
Privacy experiences → Concern	0.363	0.048	7.542	< 0.001	Significant
Privacy experiences → Risk	0.266	0.045	5.862	< 0.001	Significant
Awareness → Concern	0.663	0.098	6.746	< 0.001	Significant
Awareness → Risk	0.513	0.091	5.615	< 0.001	Significant
Realism → Willingness to share	0.092	0.041	2.227	0.026	Significant
Age → Willingness to share	0.007	0.005	1.289	0.197	n/s
Education → Willingness to share	-0.05	0.044	-1.138	0.255	n/s
Employment → Willingness to share	0.05	0.048	1.041	0.298	n/s
Gender → Willingness to share	0.188	0.099	1.891	0.059	n/s
Marker → Willingness to share	-0.02	0.064	-0.312	0.755	n/s
Frequency of use → Willingness to share	0.013	0.048	0.274	0.784	n/s
Proficiency → Willingness to share	0.021	0.072	0.286	0.775	n/s
Positive experience → Willingness to share	0.084	0.052	1.608	0.108	n/s

Note. n/s = not significant, n/a = not applicable (not hypothesized); CMIN / *df* = 1.907, CFI = 0.961, NFI = 0.921, IFI = 0.912, TLI = 0.956, RMSEA = 0.043, PCLOSE = 1.000, SRMR = .0764; R^2: concern (0.220); risk (0.140); competence (0.010); benevolence/integrity (0.000); benefits (.000); willingness to share (0.51).

1.4.3. Mediation Testing

Our model proposed the APCO constructs of concern, risk, trust, and benefit as mediators. According to traditional techniques, testing for complex mediation at the same time is not an accurate approach. Instead, mediation may be accurately tested using advanced bootstrapping tests to construct confidence intervals of the mediation effects. We followed the procedures outlined in Appendix 1E to bootstrap the effects of our mediating relationships. **Table 6** shows the results of our testing.

Table 6. Bootstrapped Confidence Interval Tests for Full and Partial Mediation Model

	Mediation test (ab) (indirect effects)			Full/partial mediation test (c')			
Proposed relationship	2.5% lower bound	97.5% upper bound	Include zero?	2.5% lower bound	97.5% upper bound	Include zero?	Type of mediation relationship
HI → PC → WTS	-0.110	-0.037	No	-0.059	0.064	Yes	Full
HI → PR → WTS	-0.070	-0.013	No	-0.059	0.064	Yes	Full
HI → C → WTS	-0.036	-0.001	No	-0.059	0.064	Yes	Full[i]
HI → B/I → WTS	-0.034	0.021	Yes	-0.059	0.064	Yes	None
HI → BN → WTS	-0.012	0.044	Yes	-0.059	0.064	Yes	None

Note. B/I = benevolence and integrity; BN = benefit; C = competence; HI = human involvement; PC = privacy concern; PR = privacy risk; WTS = willingness to share voice data.

1.4.4. Additional Analysis

We investigated the contextual effects of cognitive biases and information processing, namely salience bias (H1 through H6), default trust bias, and level of effort to decipher the influence of each on how people processed the content in privacy policies and then subsequently made disclosure decisions. We performed a MANOVA using IBM SPSS v27 to determine which of the treatment effects were statistically different when they were saliently presented than when they were not. We chose a MANOVA design over a multivariate analysis of covariance (MANCOVA) because analysis of variance (ANOVA) is better at detecting a treatment effect than does analysis of covariance (ANCOVA) across classification differences caused by manipulating more than one factor simultaneously (Schneider et al., 2015).[ii]

We ran each of the four treatments separately as a fixed factor. First, our results indicate that explicitly informing people of human involvement leads to significant increases in perceived privacy concern (p = 0.010) and perceived privacy risk (p = 0.012) and significant decreases to willingness to share (p = 0.036) (H1, H2, and H6 are supported; **Table 7**). Our analysis shows no significant effects of overt human involvement disclosure on competence trust, benevolence and integrity trust, and perceived benefits, and thus H3–H5 are not supported. Second, informing people of the device product type leads to significant increases in perceived privacy concern (p = 0.005), perceived privacy risk (p = 0.007) and significant decreases in competence trust (p =

0.000), and benevolence and integrity trust (p = 0.040) (**Table 8**). However, product type does not significantly influence willingness to share voice data, and the significant difference in perceived benefits is between the product types (i.e., Amazon Echo vs. Alibaba Genie). Third, the disclosure of cloud service provider of the product type only significantly decreases benevolence and integrity trust (p = 0.025; **Table 9**). Finally, the level of aid provided to people to increase the interpretability of the privacy policy and thus their information processing only significantly increases privacy concern (p = 0.039; **Table 10**).

Table 7. Results of Human Involvement Treatment Effects (n = 499; df = 496)

	Treatment			MANOVA Results		
Dependent Variables	Overt Mean (SD)	Covert Mean (SD)	Control Mean (SD)	p-value	Partial η2	Hypothesis supported?
Privacy concern (H1)	5.35 (1.47)	5.09 (1.45)	4.68 (1.60)	0.010	0.018	Yes
Privacy risk (H2)	4.84 (1.57)	4.46 (1.51)	4.32 (1.46)	0.012	0.018	Yes
Competence (H3)	5.21 (1.17)	5.09 (1.18)	5.36 (1.13)	0.262	0.005	No
Benevolence/integrity (H4)	4.07 (1.38)	3.95 (1.36)	4.43 (1.22)	0.074	0.010	No
Benefits (H5)	5.21 (1.19)	5.23 (1.17)	5.33 (0.95)	0.799	0.001	No
Willingness to share (H6)	3.63 (1.84)	3.93 (1.81)	4.30 (1.55)	0.036	0.013	Yes

Note. The relationship between *human involvement* and *willingness to share* (H6) is fully mediated through privacy concern (H6a) and privacy risk (H6b).

Table 8. Results of Product Treatment Effects (n = 499; df = 496)

	Treatment			MANOVA Results		
Dependent Variables	Amazon Echo Mean (SD)	Alibaba Genie Mean (SD)	Control Mean (SD)	p-value	Partial η2	Sig.?
Privacy concern	5.06 (1.48)	5.37 (1.43)	4.68 (1.60)	0.005	0.021	Yes
Privacy risk	4.43 (1.44)	4.85 (1.63)	4.32 (1.46)	0.007	0.020	Yes
Competence	5.37 (1.06)	4.95 (1.24)	5.36 (1.13)	0.000	0.032	Yes
Benevolence/integrity	4.10 (1.31)	3.92 (1.43)	4.43 (1.22)	0.040	0.013	Yes
Benefits	5.46 (1.11)	5.00 (1.20)	5.33 (0.95)	0.000	0.036	Yes
Willingness to share	3.88 (1.79)	3.70 (1.86)	4.30 (1.55)	0.095	0.009	No

Note. Although the p-value for *benefits* is significant, the hypothesis is not supported because the significant relationship is between Amazon Echo and Alibaba Genie.

Table 9. Results of Cloud Provider Treatment Effects (n = 499; df = 496)

	Treatment			MANOVA Results		
Dependent Variables	AWS Mean (SD)	Alibaba Cloud Mean (SD)	Control Mean (SD)	p-value	Partial η2	Sig.?
Privacy concern	5.23 (1.42)	5.21 (1.51)	4.68 (1.60)	0.055	0.012	No
Privacy risk	4.67 (1.54)	4.62 (1.57)	4.32 (1.46)	0.353	0.004	No
Competence	5.16 (1.19)	5.14 (1.15)	5.36 (1.13)	0.483	0.003	No
Benevolence/integrity	3.90 (1.34)	4.13 (1.40)	4.43 (1.22)	0.025	0.015	Yes
Benefits	5.22 (1.17)	5.22 (1.19)	5.33 (0.95)	0.819	0.001	No
Willingness to share	3.78 (1.86)	3.78 (1.80)	4.30 (1.55)	0.167	0.007	No

Table 10. Results of Level of Aid Treatment Effects (n = 499; df = 496)

	Treatment			MANOVA Results		
Dependent Variables	High Mean (SD)	Low Mean (SD)	Control Mean (SD)	p-value	Partial η2	Sig.?
Privacy concern	5.27 (1.46)	5.16 (1.47)	4.68 (1.60)	0.039	0.013	Yes
Privacy risk	4.71 (1.53)	4.57 (1.58)	4.32 (1.46)	0.242	0.006	No
Competence	5.11 (1.16)	5.19 (1.19)	5.36 (1.13)	0.374	0.004	No
Benevolence/integrity	3.96 (1.36)	4.07 (1.38)	4.43 (1.22)	0.079	0.01	No
Benefits	5.25 (1.04)	5.20 (1.32)	5.33 (0.95)	0.738	0.001	No
Willingness to share	3.69 (1.82)	3.89 (1.84)	4.30 (1.55)	0.081	0.01	No

1.4.5. Qualitative Analysis

We performed a qualitative analysis as a robustness check for the treatment effects and to investigate whether perceived human involvement is a salient consideration that acts as an inhibitor of consumers' willingness to share their voice data. As part of the post experiment questionnaire, we asked participants to indicate any reasons they have for why they may or may not be willing to share their voice data with companies to improve their voice services. We collected 498 validated, open-ended responses from participants (one participant did not provide any reasons for or against sharing). We used Atlas.ti v9 to analyze the responses by first examining a sample of responses to generate a preliminary list of potential reasons. We open-coded the responses, then performed axial coding of the codes to organize them based on the distinct connections of the codes. We then aggregated the axial codes into six main categories based on the constructs in the research model (i.e., concerns, risks, trust, benefits, human involvement, and other) for sharing or not sharing voice data. Each category contained multiple subcategories. For example, the concern category included opposite-end responses where some participants expressed high concerns, whereas other

participants did not have any concerns. Participants confirmed human involvement as a specific privacy-related concern as illustrated by the following quotes:

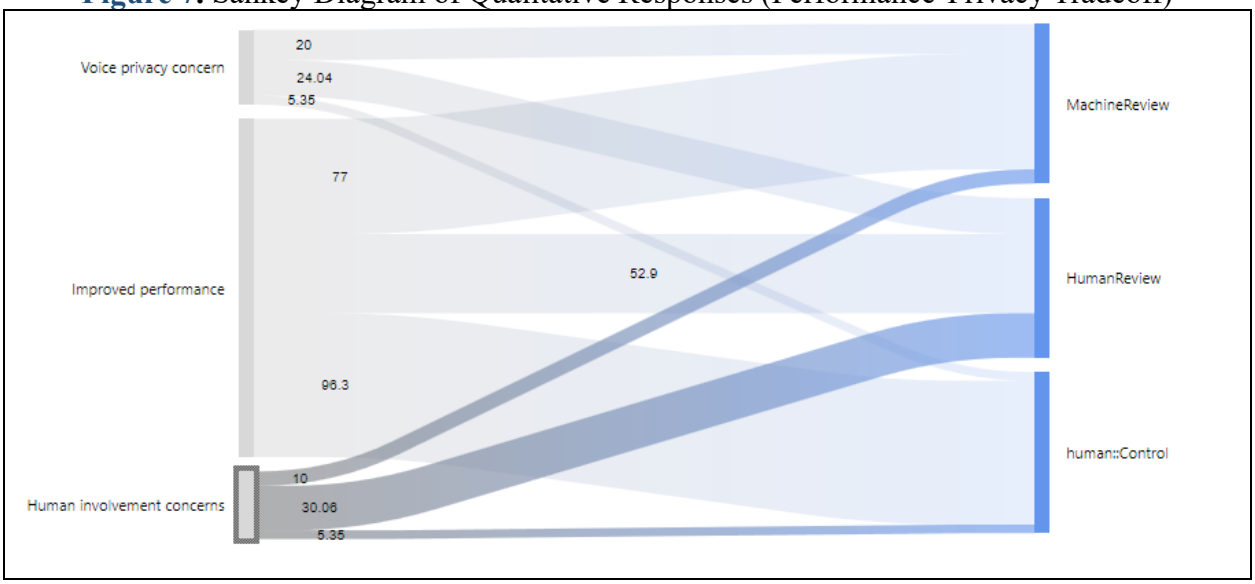
> "It seems sketchy to me that the voice data might be shared with humans. It seems like an invasion of privacy." (Respondent 230, human reviewer group)

> "This assistant is in my home hearing all kinds of conversations and I would not like to share all of this with others. It seems way too open to others to me." (Respondent 466, machine reviewer group)

> "The idea of having humans listening to my voice recordings, even in the context of making the service better, is really unsettling. I realize on some level that this is probably standard across all personal assistant devices, but actually reading the privacy policy and thinking back to conversations I've had around my personal assistant devices makes me concerned for what exactly is out there of my personal information." (Respondent 84, human reviewer group).

Figure 7 depicts a Sankey diagram of coded open responses from participants in the (1) machine reviewer, (2) human reviewer, and (3) control (no mention of reviewer) conditions about their (a) voice privacy concerns, (b) human involvement concerns, and (c) interests to improve the performance of the voice systems. Participants in the human reviewer condition expressed human involvement and voice privacy concerns at higher levels than participants in the other conditions. Participants in the human reviewer condition also expressed lower interest to improve voice system performance than the control and machine reviewer conditions. The qualitative responses corroborated the quantitative results, such that participants who were cognizant of human involvement expressed this as a salient concern. Further, their privacy concern exceeded the improved performance benefit, which shifted their performance-privacy tradeoff calculus toward preserving privacy. Participant quotes are included in Appendix 1F.

Figure 7. Sankey Diagram of Qualitative Responses (Performance-Privacy Tradeoff)



1.5. Discussion

The ability of agentic IS artifacts to autonomously perceive their environments and collect consumers' sensitive data has catalyzed ongoing debates regarding regulators' role in protecting consumers through comprehensive data privacy laws like the GDPR, CCPA, and VCDPA. Research shows that human agents are heavily involved in IS and ML systems. However, when human agents are involved in ML systems, they have access to consumers' sensitive unstructured data such as image and voice data and must label and ascribe contextualized meanings to them for algorithms to interpret. Our research addresses the concern of human involvement in the degree of salience bias (i.e., overt vs. covert disclosure on companies' privacy policies) and context (i.e., cognitive biases operationalized as product type, cloud provider, and level of aid) to elicit consumers' behavioral reactions to the practice of supervised machine learning in conversational agents. To test our theorization, we conducted a qualitative and quantitative experiment that supports most of our privacy-related hypotheses. We conclude by summarizing our results and discussing their implications for theory, research, and policy.

1.5.1. Summary of Results

We tested the effects of degree of changes to company privacy policies on consumers' willingness to share their voice data. We found strong evidence for the effects of our privacy policy treatments in our experiment (see Tables 7–10). First, our baseline analysis indicates that perceived human involvement elicits higher levels of perceived privacy concern and risk (H1 and H2 supported). Human involvement also increases competence trust (H3 supported). However, human involvement does not influence benevolence and integrity trust and perceived benefits (H4 and H5 not supported). Thus, the "performance" argument posited by companies that including human reviewers is beneficial for consumers is not as important as the privacy tradeoff consumers must make to facilitate such performance improvement. Next, we found that the effect of perceived human involvement on willingness to share voice data is fully mediated only through privacy concern and privacy risk (H6a and H6b supported; H6c–e not supported). **Table 7** shows similar treatment effects of the human involvement disclosure, which corroborates the findings in the structural model. However, the effect of human involvement on competence trust is not supported as a treatment effect alone. These findings suggest that the perception of human involvement is more negatively viewed by consumers and that the benefits companies describe about involving human agents do not manifest in the decision making of consumers.

Next, we tested whether the contextual effects of cognitive biases and information processing influenced consumers' privacy and trust considerations and their willingness to subsequently share their voice data by assessing the effects of our operationalized privacy policies. This was possible because we altered key signals on the privacy policies such as company brand and the level of aid we provided to assist participants in interpreting the content of the privacy policy. Our findings show that product type has a considerable effect on participants. Namely, when participants are not given information about the product brand of the device they are using, their default trust bias

typically results in reduced privacy concerns and risk perceptions, as well as increased perceptions of competence trust, benevolence and integrity trust, and perceived benefits. (**Table 8**). Furthermore, default trust biases are magnified depending on the specific company brand that is identified (i.e., Amazon vs. Alibaba). However, we did not find support that company product type significantly influences participants' willingness to share. Also, when participants are informed of the cloud provider that specifically handles their collected data, whereas the product type is the interface that collects the data, our results show that generally participants' privacy decision making is largely unaffected (**Table 9**). We found that only benevolence and integrity trust is significantly reduced when the cloud provider is disclosed. This suggests that participants may not see a distinction between the front-end product with which they interface and the companies operating the back-end servers that process their data, which interestingly represents another form of salience bias. Finally, our results show that when participants are provided aids (comments) on the privacy policy to assist their interpreting the content, only privacy concern is significantly elevated (**Table 10**). This supports the general finding from the baseline model that participants viewed human involvement pessimistically rather than optimistically.

1.5.2. Implications for Research, Practice, and Policy

Our study yields several implications for research, practice, and policy. *First*, the study responds to calls on explainable AI (Rai, 2020) and ethical AI (De Cremer & Kasparov, 2021), both of which carry societal implications. Our results indicate the possible difficulties companies may have when they involve human agents to process consumers' collected data while also maintaining consumers' privacy rights (Grant & Wischik, 2020). The right to explanation and the right to privacy are both granted to consumers in comprehensive data privacy laws such as the GDPR, CCPA, and VCDPA; however, a fundamental assumption in the legislation is that ML systems operate similarly to traditional IS (Grant & Wischik, 2020). This is not the case because human

agents are involved in the training, explaining, and reviewing of the underlying consumer data used to train ML algorithms (Ågerfalk, 2020; Buxmann et al., 2021; Collins et al., 2021; Fügener et al., 2021; Grønsund & Aanestad, 2020; Kordzadeh & Ghasemaghaei, 2021). Human agents play a crucial role in interpreting socially meaningful cues and deducing contextual meanings from training data, which helps elucidate how conversational agents process a specific consumer's information. Our research demonstrates that human involvement is a vital factor for consumers when deciding whether to share their data to enhance a company's product. Companies may consider full machine processing of consumers' sensitive data similar to the classifier used by Apple in its Child Sexual Abuse Material (CSAM) (Apple, 2021). The automated detection system is designed to prioritize user privacy, utilizing a matching process that identifies matches without exposing the results for human review. Specifically, it employs cryptography and conducts on-device matching of hashes supplied by a governing entity, comparing these hashes with those in a database. Nevertheless, cryptography and hash matching have their limitations, as Apple's CSAM system still necessitates human intervention when a certain detection threshold is reached within a user's iCloud storage (Apple, 2021).

Second, research explains human involvement as the distinction between data privacy and information privacy (Clarke, 1997), thus, companies may consider making a concerted effort to remove company employees or third parties from the training data process altogether and allow consumers to *train* their own data. Functionality may be granted directly to consumers to train their own ML algorithms for personalization services. For example, in the iOS 15, Apple's Siri can work without an internet connection by using on-device speech recognition whereby the audio never leaves a device and the data are neither sent nor stored on the cloud for processing (Vincent, 2021). Similar to the way consumers can access on-device privacy and security settings, they could

access on-device machine learning features settings to tailor and tweak the functionality of their on-device machine learning algorithms to personalize them to their desired preferences (Vincent, 2017).

Third, for policymakers, this study establishes that regulations should reflect the difficulty of involving human agents to detect algorithmic biases and explain an algorithm's actions, while preserving the data privacy of consumers by which the algorithms were trained. Thus, bridging the gap between technology and law may address the conflict between privacy and explainability in the law (Villaronga et al., 2018). Currently, researchers and developers implement privacy-enhancing technologies to anonymize and obscure consumers' data. Legislation may be adopted that revises the concept of explainability in light of considering techniques such as differential privacy and hashing of sensitive data. However, because technologies used to obscure training data also affect the ability to detect biases and systematic unfairness in ML systems, human-ML augmentation may remain the favored practice (Teodorescu et al., 2021).

Fourth, using a privacy lens, we extend discussion in the human-ML augmentation discourse by conceptualizing a new construct, *perceived human involvement*, as a proposed antecedent to the APCO model, one of which can predict outcomes related to consumers' willingness to share training data and to opt out of training data participation. A key implication of this finding is that researchers and practitioners who investigate human-ML dynamics should strongly consider privacy concerns, because new legislation may empower consumers to restrict the sharing of their training data and thus inhibit the sufficient development of ML systems. For example, large language models are used to train generative AI chatbots, such as ChatGPT and Bing AI, and usually have a billion or more parameters to infer new content. Over time, if legislation is passed that allows consumers to restrict the use of their data in training generative AI systems, then

companies may have to rely on old data sets that may not provide personalized inferences that are readily useful to the consumers. The need to protect consumers' privacy can stifle the improvement of generative AI systems, which creates another conundrum beyond the scope of this study. This implication extends to discussions on algorithmic biases, because the ability to obtain and process data sets with complete and diverse features is necessary to reduce algorithmic biases and systematic unfairness in the decisions and outputs rendered by these systems.

1.5.3. Limitations and Future Research

One of this study's limitations is that the proposed experiment was performed in the context of conversational ML systems with only two types of virtual assistants, Amazon Alexa and Alibaba Genie. Future research should carefully examine the extent to which consumers experience trust and privacy concerns related to the collection and labeling of their *video* and *image* data. The labeling of images and video is exceedingly important, as discrimination and biases in computer vision ML systems can be vastly perpetuated and easily observed. For example, Facebook experienced a data labeling problem in its AI system and mistakenly labeled Black men in altercations with white civilians and police officers as "Primates" (Mac, 2021). Specifically, the AI recommendation system asked Facebook users if they would like to "keep seeing videos about Primates." To overcome racism and discrimination in image recognition software, human involvement and data annotation may be strongly recommended, and thus consumers may perform a different calculus when deciding to sacrifice their privacy for the dramatic enhancement in the performance of facial recognition technologies.

A second limitation is on our choice of method. We employed a scenarios-based survey with a 2 x 2 x 2 x 2 factorial design and forced a single control condition across all the treatments. Although we used this single control condition as a comparison to determine effects of the factors, we could not identify the precise effects of individual factors (Vance et al., 2015). Because this is

the first study to report evidence of human involvement concerns, we encourage future research to corroborate these findings with evidence from the factorial survey method. Notably, researchers can build full factorial designs that include control conditions for each factor. All possible combinations of each factor at its different levels form a Cartesian product and ensure an orthogonal design, which avoids ordering bias and reduces multicollinearity to nearly zero (Jasso, 2006; Vance et al., 2015). Orthogonality will allow researchers to clearly distinguish the effects among the different factors presented on a vignette treatment.

A third limitation pertains to the generalizability of our results to citizens in other countries. The APCO includes culture as an antecedent to privacy concern (Dinev et al., 2015; Smith et al., 2011). Our study includes responses from only US participants, and only two states (California and Virginia) at present have comprehensive data privacy laws. Therefore, US participants may not feel as strongly about privacy as citizens in the European Union, for example, who are granted numerous data privacy rights by the GDPR. Thus, future research can test the robustness of the proposed model in a more generalizable manner by examining the trust and privacy considerations of participants in the EU. Conversely, because we contrasted American-based products with Chinese-based products and surveyed a US population, future research may present a similar study to a Chinese population to understand participants' perceptions of trust and privacy toward human involvement. Because privacy laws differ in China, the EU, and the US, future studies can yield contextual nuances that can better inform the discourses on privacy and human-ML augmentation.

1.6. Conclusion

The discourse on human-ML augmentation is predominantly situated within organizational and societal contexts. It appears that because of a heavy focus on understanding the benefits and consequences of large-scale deployments of ML systems, IS research has overlooked the privacy costs to consumers when organizations augment ML systems with human agents to improve their

systems. We use the enhanced APCO model to investigate whether perceived human involvement may contribute to consumers' privacy calculus when deciding whether to share their voice data with companies. Because we found evidence to support these hypotheses, this study introduces perceived human involvement as a new construct to the information privacy discourse, which can help establish a new lens toward investigating privacy concerns in ML.

1.7. Appendix 1A – Background Research and Research Context

Our literature search included an examination of journals in IS, Management, and Marketing, namely special issues on "artificial intelligence," for articles that described the level of human involvement or the roles of human agents in artificial intelligence, machine learning, automation, and augmentation systems. We found numerous articles that explicitly referenced the roles of human agents in *training*, *explaining*, and *reviewing* data and the results generated by or inputted into machine learning systems. We described the aim of the articles, the description of the context related to the human-ML system interaction, and related concepts, constructs, and findings. The article types were mostly conceptual, followed by qualitative design. Quantitative design studies were the least observed types of articles in our literature review.

Table A1. Summary of Select References on Human Involvement in ML Systems

Activity	Purpose of activity	Key references on human involvement in ML systems
Explainability is the obligation expressed in the GDPR (Parliament and Council of the European Union, 2016) requiring data controllers that handle and process data to supply users with "meaningful information about the logic involved in automated decision-making" (Grant & Wischik, 2020, p. 1352). As a result, a new class of systems was developed, explainable AI, to explain and provide visibility into its decision-making logic and to provide reliable indications of its future behaviors (Rai, 2020)	• Detect bias in the dataset or output • Monitor and review algorithmic output • Audit and alter the ML model(s) • Interpret fairness and ethics of decisions rendered by ML systems • Analyze the logic of ML systems logic in decision-making • Retain a human-in-the-loop to augment the ML system	Ågerfalk (2020): Algorithms can process decontextualized behavioral data, but humans are needed to interpret the socially meaningful signs from the context in which the behaviors are situated.
		Asatiani et al. (2021): GDPR's provision for a right to explanation could prompt data handlers to provide meaningless explanations that are not easily interpretable by the affected data subject, prompting calls for human-centered approaches to explainability.
		Benbya et al. (2021) Managers assume removing humans from the loop reduces human biases; however, biases exist in training datasets, noisy data, and statistical error, requiring that humans remain in the loop.
		Ge et al. (2021) Humans can identify possible failure points in predictive models and can reduce the time to build deployable ML models.
		Gregory et al. (2020): Platform users will perceive greater value if platforms increase the level of explainability in the predictions made by their ML systems.
		Grønsund and Aanestad (2020): A human-in-the-loop pattern yields the best outcome for organizations deploying ML systems, whereby humans continually augment the algorithm through auditing and altering the algorithm.
		Kellogg et al. (2019): Algorithms can be opaque and difficult to decipher, to a point where humans are unable to understand or interpret the models.
		Marabelli et al. (2021): Humans may be required to systematically vet or to identify problematic

	• Remove algorithm opacity and increase transparency	decisions made by algorithms.
		Markus (2017) Explaining decisions from ML systems with only minimal human intervention is nearly impossible.
		Raisch and Krakowski (2021): With unsupervised learning models, managers deduce patterns and themes from unlabeled data, of which they were previously unaware.
		Teodorescu et al. (2021): ML tools and humans cannot operate independently to achieve fairness in decision-making. Only through their joint decision-making may it be possible.
Right to erasure ("right to be forgotten") is the right of individuals to request to have their personal data collected by others to be erased and no longer processed. Deleting data from an ML model has given rise to the notion of machine unlearning.	• Regulatory or legal compliance • Remove data subject's personal data from training data sources, model weights, and parameters • Retrain ML models	Bourtoule et al. (2021): ML models memorize training data and to make models forget requires knowing how individual training points contributed to the original model parameters. Deleting the training data and retraining the model from scratch incurs large computational and time overhead. Sharded, Isolated, Sliced, and Aggregated (SISA) training is proposed as a method for machine unlearning.
		Cao and Yang (2015): Machine unlearning is the process for making ML systems forgot. The objective is to remove targeted training data and revert the models' effects to operate as "if the data had never existed" (p. 464).
		Chen et al. (2021a): Machine unlearning generates two ML model versions: the *original* and the *unlearned* models. Unintended privacy risks may occur when imprints are left in the unlearned model and adversaries perform membership inference attacks.
		Villaronga et al. (2018): Impossible to fully comply with the Right to be Forgotten in AI environments because of the impractical nature of deleting data across shared environments and networks.
Data annotation is the labeling and categorizing of data for machine learning model training and AI applications (Appen, 2021). ML applications rely on labeled input-output pairs to classify or categorize new data inputs into known, labeled output classes, a process known as supervised learning (Lebovitz et al., 2021).	• Train and retrain ML model	Benbya et al. (2021): In human-machine augmentation, machines learn from humans through training datasets and humans learn from machines through the patterns detected in the machine output.
		Borges et al. (2021): Classic AI, classic ML, and deep learning algorithms depend on human training, programming, or knowledge.
		Fügener et al. (2021): AI is trained by humans
		Kane et al. (2021): ML systems are designed using training data containing human biases rather than through logic-based design built with code.
		Lebovitz et al. (2021): Qualified experts provide ground truth labels to train and validate ML models in a US hospital setting.
		Raisch and Krakowski (2021): Supervised learning begins with human domain expertise where managers provide labeled training data to machines that analyze the data to develop models and generate rules for algorithmic actions and decision-making.
		This study: We investigate the effects of perceived human involvement in data annotation practices through a privacy and trust lens to empirically understand the "privacy" side of the privacy-explainability paradox.

1.8. Appendix 1B – Survey and Measurement Details

Table B1. Demographics and Smart Device Use

Variable	Item	Source
	Prompt: Please know that we are uninterested in your identity and collect demographic data to report only aggregate level results.	
Gender	Please indicate your identified gender: [Male / Female / Prefer not to say / Other [please specify]]	Adapted from Al-Natour et al. (2020)
Age	Please indicate your age in years: (Must be 18 or above)	Adapted from Al-Natour et al. (2020)
Education	What is the highest level of education you have completed? 1 = Less than high school / secondary school 2 = High school / secondary school 3 = Some university, but have not completed a degree 4 = Associate degree 5 = Bachelor's degree 6 = Master's degree 7 = Doctorate / Ph.D. [Radio button for each selection]	Adapted from Al-Natour et al. (2020)
Annual income range	Approximately, what is your annual income range? 1 = < $30,000 2 = $30,001 – $75,000 3 = $75,001 – $150,000 4 = $150,001 – $300,000 5 = $300,001 – $500,000 6 = $500,001+ [Radio button for each selection]	Adapted from Al-Natour et al. (2020)
Employment status	Please indicate your current employment status: 1 = Employed part-time 2 = Employed full-time 3 = Not employed 4 = Self-employed 5 = Student 6 = Retired 7 = Other [Radio button for each selection]	Adapted from Al-Natour et al. (2020)
Ethnicity	Please indicate the ethnic group you most identify with: 1 = American Indian or Alaskan Native 2 = Asian 3 = Black or African American	Adapted from Crossler and Bélanger (2019)

	4 = Hispanic or Latino 5 = Middle Eastern or North African 6 = White or Caucasian 7 = Prefer not to say [Radio button for each selection]	
Familiarity with voice assistants	Are you familiar with virtual assistants (i.e., voice recognition systems) such as Alibaba AliGenie, Apple Siri, Amazon Alexa, or Google Assistant? [Yes / No]	Screening question
Frequency of use of voice assistants	How often do you use a virtual assistant (e.g., Alibaba AliGenie, Apple Siri, Amazon Alexa, Google Assistant, and the like)? 1 = Never 2 = Less than once a month 3 = 1–3 times a week 4 = Once a day 5 = 2–4 times a day 6 = More than 4 times a day [Radio button for each selection]	Adapted from Lopatovska et al. (2018)
Smartphone proficiency	How would you evaluate your smartphone skills in general? [1 = poor, 2 = fair, 3 = good, 4 = very good, 5 = excellent]	
Mobile OS	Indicate the mobile operating system installed on your mobile phone [1 = iOS, 2 = Android, 3 = Windows Phone, 4 = Other]	Adapted from Al-Natour et al. (2020)
Mobile OS update	Have you updated your mobile phone operating system during the last 12 months? [Yes / No]	
Mobile phone ownership (months)	How long have you owned a smart device such as a smartphone or smart speaker? 1 = Less than 12 months 2 = 13–24 months 3 = 25–36 months 4 = 37–48 months 5 = 49–60 months 6 = 61–72 months 7 = More than 72 months [Radio button for each selection]	Adapted from Malhotra et al. (2004)

Table B2. Pre-experiment Survey Control Variables

Variable	Contextualized items	Original items	Source
Disposition to trust: benevolence	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with people, in general:" 1. Same 2. Same	1. In general, people really do care about the well-being of others 2. The typical person is sincerely concerned about the problems of others	Retained from Moody et al. (2014)

	3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	3. Most of the time, people care enough to try to be helpful, rather than just looking out for themselves	
Disposition to trust: competence	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with people, in general:" 1. Same 2. Same **(Reverse coded)** A large majority of professional people are **incompetent** in their area of expertise [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I believe that most professional people do a very good job at their work 2. Most professionals are very knowledgeable in their chosen field 3. A large majority of professional people are competent in their area of expertise	Retained from Moody et al. (2014)
Disposition to trust: integrity	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with people, in general:" 1. Same 2. Same 3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. In general, most folks keep their promises 2. I think people generally try to back up their words with their actions 3. Most people are honest in their dealings with others	Retained from Moody et al. (2014)
Response set item	1. **Prompt**: "If 2 + 3 = 5, then select "Disagree" as the response to this question."	N/A	
Disposition to distrust: malevolence	**Prompt**: "Read carefully and indicate your general agreement with each of the following statements about online merchants:" 1. Same 2. Same 3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I worry that online merchants are usually concerned about their own good 2. It concerns me a lot that online merchants pretend to care more about their customers than they really do 3. I fear that most online merchants inwardly dislike putting themselves out to help out their customers	Retained from Moody et al. (2014)
Disposition to distrust: incompetence	**Prompt**: "Read carefully and indicate your general agreement with each of the following statements about online merchants:" 1. Same 2. Same 3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I am troubled that many online merchants are not as knowledgeable in their product/service area as you would expect 2. I am cautious because I believe that most online merchants do a haphazard job at what they do 3. Concern is justified, since many online merchants are not really competent in their area of expertise	Retained from Moody et al. (2014)
Disposition to distrust:	**Prompt**: "Read carefully and indicate your general agreement with each of the following statements about online merchants:"	1. Unfortunately, most online merchants would tell a lie if they could gain by it	Retained from Moody et al. (2014)

deceit	1. Same 2. Same 3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	2. It's a troubling fact that online merchants don't always hold to the standard of honesty they claim 3. Sadly, most online merchants would cheat their customers if they thought they could get away with it	
Blue attitude marker variable	**Prompt**: "Indicate your color preference. I . . ." 1. ". . . prefer blue to other colors" 2. ". . . like the color blue" 3. ". . . like blue clothes" [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I prefer blue to other colors 2. I like the color blue 3. I like blue clothes	Miller and Chiodo (2008)

Table B3. Attention Check Questions During Experiment

Variable	Item	Description
Default trust bias: product type	**Prompt:** The brand of the smart speaker is clearly: 1. Apple HomePod 2. Google Nest 3. Amazon Echo 4. Alibaba Genie 5. Baidu Xiaodu 6. Samsung Galaxy 7. Company brand not listed [Radio button for each selection: End of survey for invalid responses]	Attention check #1: Product purchased
Default trust bias: cloud service provider	**Prompt**: The cloud storage company hosting the data collected by the smart speaker is clearly: 1. Apple iCloud 2. Google Cloud 3. Amazon Web Services 4. Alibaba Cloud 5. Baidu AI Cloud 6. Samsung AppStack 7. Company not listed [Radio button for each selection: End of survey for invalid responses]	Attention check #2: Cloud storage provider
Information processing: level of aid provided	**Prompt**: For the privacy policy you reviewed, which types of aids were provided to help you understand the content: 1. No aids were provided, only the privacy policy content 2. Callout boxes only to highlight relevant passages 3. Callout boxes and explanations of the highlighted passages 4. I did not see a privacy policy. It was not provided. [Radio button for each selection: End of survey for invalid responses]	Attention check #3: Level of aid provided to assist cognitive processing
Salience bias:	**Prompt**: The privacy policy clearly states which type of actor will assist in the supervised machine learning training:	Attention check

human involvement	1. Humans to manually listen to voice samples to help [Alexa/AliGenie/your smart device] understand correct interpretations 2. [Alexa/AliGenie/Your smart device] to collect voice samples to help it understand correct interpretations 3. The privacy policy did not state anything about supervised machine learning training 4. I did not see a privacy policy. It was not provided. [Radio button for each selection: End of survey for invalid responses]	#4: Explicit mention of human reviewer (overt vs covert)

Table B4. Post Experiment Survey Items

Variable	Contextualized items	Original items	Source
Realism item	**Prompt**: "Read carefully and indicate your agreement with the following statements about information sharing . . ." 1. I could imagine my voice data being shared with multiple parties **[Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]**	**I could imagine a similar scenario taking place at my company**	**Adapted from Barlow et al. (2018)**
Willingness to share voice data for processing	Prompt: "Suppose [Amazon/Alibaba] provided a privacy setting option that allowed you to opt in or opt out of sharing your voice data with the company to improve its voice recognition service: [Alexa/Alibaba Genie]. With this option, please answer the following: "I would be willing to share my voice data with [Amazon/Alibaba]. . ." 1. . . . to improve its voice recognition service 2. . . . for data analysis purposes 3. . . . to receive personalized services [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	''Suppose you wanted more specific information about a given product and you could consult (one time only) by telephone with a salesman from the seller for 15–30 min (free of charge). For this service, please answer the following:'' 1. I would be willing to provide information like my name, address and phone number to the seller's representative 2. I would be willing to provide my social security number to the seller's representative 3. I would be willing to share the specifics of my product needs with the seller's representative	Adapted from Moody et al. (2017)
(Open-ended) Willingness to share voice data for processing	Please indicate any reasons why or why not you would be willing to share your voice data with [**Amazon/Alibaba**]:	N/A	N/A
Cognitive processing load (manipulation check for level of effort (aid) variable)	**Prompt**: "Read carefully and indicate to which extent each of the following statements describes you when reading through the privacy policy content:" 1. It generally took me a lot of processing efforts to figure out **how to interpret the content in the privacy policy** **2.** I needed a lot of thinking when deciding how to **interpret the privacy policy**	1. In the study, it generally took me a lot of processing efforts to figure out **how to find a target page/content on the Web site** 2. I needed a lot of thinking when deciding how to navigate from a current page toward the target page/content on the Web site 3. In general, I spent a lot of cognitive effort to **find a**	Adapted from Fang et al. (2012)

	3. In general, I spent a lot of cognitive effort to **understand the content in the privacy policy** 4. Generally speaking, my **reading the privacy policy** was cognitively demanding 5. Overall, I incurred a significant cognitive load when trying to **understand the privacy policy content** [Likert-type 7-point scale: 1 = Describes me very poorly: 7 = Describes me extremely well]	**target page/content on the Web site** 4. Generally speaking, my navigating the Web site to locate a target page/ content was cognitively demanding 5. Overall, I incurred a significant cognitive load when trying to **find a target page/content on the Web site**	
Perceived human involvement (manipulation check for (c)overt privacy policy)	**Prompt**: "Read carefully and indicate the level of involvement you believe human reviewers have in evaluating the service of **[Alexa/Alibaba Genie].** 1. To what extent do you believe **human reviewers** are involved in the **training of [Alexa/Alibaba Genie]?** 2. Please indicate how closely human reviewers work with your voice data to improve [Alexa/Alibaba Genie]. [Likert-type 7-point scale: 1 = Never; 7 = Always] **Prompt**: "Read carefully and indicate your agreement with each of the following statements about human reviewers:" 1. I feel that human reviewers are involved in the training of [Alexa/Alibaba Genie]. 2. The majority of training to improve [Alexa/Alibaba Genie] skills require human reviewer involvement. 3. In the training of [Alexa/Alibaba Genie], human reviewers are asked to give their opinions. [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. To what extent have **you** been involved in the **design of the IRIS-SAP-FI system?** (Sasidharan et al., 2011) 2. To measure stakeholder involvement through business analysts (averaged across all groups.) For each architecture team, please indicate how closely **each architecture team** works with **business analysts**. (Scale: Business analysts (1) work within . . .; (2) work closely with . . .; work occasionally with . . .; (4) do not work with . . . the architecture group) (Boh & Yellin, 2006) 3. I feel that **the employees** have been involved in the **Food at Work project** (Lassen et al., 2007) 4. The majority of the **ward team was involved in the decision to use the C-MIS** (Segaar et al., 2007) 5. In the **decision-making process about the C-MIS I** was asked to give **my** opinion (Segaar et al., 2007)	See item for source
Trusting beliefs: benevolence	**Prompt**: "Read carefully and indicate the extent to which you believe each of the following statements about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. I believe that **[Amazon/Alibaba]** would act in my best interest 2. If I required help, **[Amazon/Alibaba]** would do its best to help me 3. **[Amazon/Alibaba]** is interested in my well-being, not just its own	1. I believe that **the seller** would act in my best interest 2. If I required help, **the seller** would do his or her best to help me 3. **The seller** is interested in my well-being, not just his or her own	Adapted from Moody et al. (2017)

	[Likert-type 7-point scale: 1 = Extremely unbelievable: 7 = Extremely believable]		
Trusting beliefs: competence	**Prompt**: "Read carefully and indicate the extent to which you believe each of the following statements about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. **[Amazon/Alibaba]** would be competent and effective in providing **its voice recognition service**. 2. **[Amazon/Alibaba]** would perform its role of providing opportunities for **its voice recognition service** very well. 3. Overall, **[Amazon/Alibaba]** would be a capable and proficient provider of **its voice recognition service**. 4. In general, **[Amazon/Alibaba]** would be very knowledgeable about **its voice recognition service**. [Likert-type 7-point scale: 1 = Extremely unbelievable: 7 = Extremely believable]	1. **The seller** would be competent and effective in providing the **product** 2. **The seller** would perform his or her role of providing opportunities for the **product** very well 3. Overall, **the seller** would be a capable and proficient provider of the **product** 4. In general, **the seller** would be very knowledgeable about the **product**	Adapted from Moody et al. (2017)
Trusting beliefs: integrity	**Prompt**: "Read carefully and indicate the extent to which you believe each of the following statements about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. **[Amazon/Alibaba]** would be truthful in its dealings with me 2. I would characterize **[Amazon/Alibaba]** as honest 3. **[Amazon/Alibaba]** would keep its commitments 4. **[Amazon/Alibaba]** would be sincere and genuine [Likert-type 7-point scale: 1 = Extremely unbelievable: 7 = Extremely believable]	1. **The seller** would be truthful in his or her dealings with me 2. I would characterize **the seller** as honest 3. **The seller** would keep his or her commitments 4. **The seller** would be sincere and genuine	Adapted from Moody et al. (2017)
Distrusting beliefs: malevolence	**Prompt**: "Read carefully and indicate to which extent each of the following statements describes your feelings about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. I worry that **[Amazon/Alibaba]** is only concerned about its own interests 2. It concerns me a lot that **[Amazon/Alibaba]** pretends to care more about me than it really does 3. I fear that **[Amazon/Alibaba]** inwardly dislikes putting itself out to help other customers [Likert-type 7-point scale: 1 = Does not describe my feelings: 7	1. I worry that **the seller** is only concerned about his or her own interests 2. It concerns me a lot that **the seller** pretends to care more about me than he or she really does 3. I fear that **the seller** inwardly dislikes putting himself or herself out to help other buyers	Adapted from Moody et al. (2017)

	= Completely describes my feelings]		
Distrusting beliefs: incompetence	**Prompt**: "Read carefully and indicate to which extent each of the following statements describes your feelings about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. I am troubled that **[Amazon/Alibaba]** is not as knowledgeable in its field as I would expect 2. I am cautious because I believe that **[Amazon/Alibaba]** does a haphazard job at what it does 3. Concern is justified, since **[Amazon/Alibaba]** is not really competent in its area of expertise [Likert-type 7-point scale: 1 = Does not describe my feelings: 7 = Completely describes my feelings]	1. I am troubled that **the seller** is not as knowledgeable in his or her field as I would expect 2. I am cautious because I believe that **the seller** does a haphazard job at what he or she does 3. Concern is justified, since **the seller** is not really competent in his or her area of expertise	Adapted from Moody et al. (2017)
Distrusting beliefs: deceit	**Prompt**: "Read carefully and indicate to which extent each of the following statements describes your feelings about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. Unfortunately, **[Amazon/Alibaba]** would tell a lie if it could gain by it 2. It's a troubling fact that **[Amazon/Alibaba]** won't always hold to the standard of honesty it claims 3. Sadly, **[Amazon/Alibaba]** would cheat on its financial statements if it thought it could get away with it [Likert-type 7-point scale: 1 = Does not describe my feelings: 7 = Completely describes my feelings]	1. Unfortunately, **the seller** would tell a lie if he or she could gain by it. 2. It's a troubling fact that **the seller** won't always hold to the standard of honesty he or she claims 3. Sadly, **the seller** would cheat on his or her financial statements if he or she thought he or she could get away with it	Adapted from Moody et al. (2017)
Response set item	1. **Prompt**: "If there are 7 days in a week, then select "Neither believable nor unbelievable" as the response to this question."	N/A	
Privacy concern: perceived surveillance	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Alexa/AliGenie]** (the smart speaker) . . ." 1. I believe that the **audio captured by [Alexa/AliGenie]** is monitored at least part of the time 2. I am concerned that **[Alexa/AliGenie]** is collecting too much information about me 3. I am concerned that **[Alexa/AliGenie]** may monitor my personal activities [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly	1. I believe that the **location of my mobile device** is monitored at least part of the time. 2. I am concerned that **mobile apps** are collecting too much information about me. 3. I am concerned that **mobile apps** may monitor my activities on my mobile device.	Adapted from Xu et al. (2012a)

	agree]		
Privacy concern: perceived intrusion	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Alexa/AliGenie]** (the smart speaker) . . ." 1. I feel that as a result of my **interacting with [Alexa/AliGenie]**, others know about me more than I am comfortable with 2. I believe that as a result of my **interacting with [Alexa/AliGenie]**, information about me that I consider private is now more readily available to others than I would want 3. I feel that as a result of my **interacting with [Alexa/AliGenie]**, information about me is out there that, if used, will invade my privacy [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I feel that as a result of my **using mobile apps**, others know about me more than I am comfortable with. 2. I believe that as a result of my **using mobile apps**, information about me that I consider private is now more readily available to others than I would want. 3. I feel that as a result of my **using mobile apps**, information about me is out there that, if used, will invade my privacy.	Adapted from Xu et al. (2008) and Xu et al. (2012a)
Privacy concern: secondary use of personal information	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Alexa/AliGenie]** (the smart speaker) . . ." 1. I am concerned that **[Alexa/AliGenie]** may use my personal information for other purposes without notifying me or getting my authorization 2. When I give personal information to **[Alexa/AliGenie]**, I am concerned that it may use my information for other purposes. 3. I am concerned that **[Alexa/AliGenie]** may share my personal information with other entities without getting my authorization. [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. I am concerned that **mobile apps** may use my personal information for other purposes without notifying me or getting my authorization. 2. When I give personal information to use **mobile apps**, I am concerned that apps may use my information for other purposes. 3. I am concerned that **mobile apps** may share my personal information with other entities without getting my authorization.	Adapted from Smith et al. (1996) and Xu et al. (2012a)
Response set item	**Prompt**: "If there is one T in the word "Tech," then select "Strongly agree" as the response to this question."	N/A	

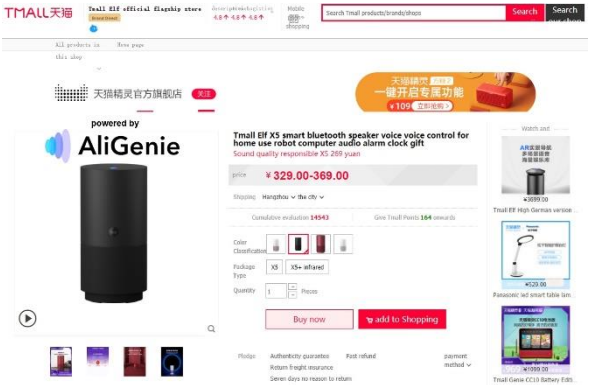
Perceived privacy risks of voice systems	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Alexa/AliGenie]** (the smart speaker) . . ." 1. In general, it would be risky to give my personal information to **[Alexa/AliGenie]**. 2. There would be high potential for loss associated with giving my personal information to **[Alexa/AliGenie]**. 3. There would be too much uncertainty associated with giving my personal information to **[Alexa/AliGenie]**. 4. Providing **[Alexa/AliGenie]** with my personal information would involve many unexpected problems [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. In general, it would be risky to give my personal information to **commercial/government websites**. 2. There would be high potential for loss associated with giving my personal information to **commercial/government websites**. 3. There would be too much uncertainty associated with giving my personal information to **commercial/government websites**. 4. Providing **commercial/government websites** with my personal information would involve many unexpected problems	Adapted from Hong and Thong (2013)
Perceived benefits of [Alexa / AliGenie]	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Alexa/AliGenie]** (the smart speaker) . . ." 1. **[Alexa/AliGenie]** reduces my searching time to find the information that I need. 2. **[Alexa/AliGenie]** can provide me with the convenience to instantly access the information that I need. 3. Overall, I feel that using **[Alexa/AliGenie]** is beneficial. [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. **M-Coupon service** reduces my searching time to find the promotional information that I need. 2. **M-Coupon service** can provide me with the convenience to instantly access the promotional information that I need. 3. Overall, I feel that using **M-Coupon service** is beneficial.	Adapted from Xu et al. (2011b)

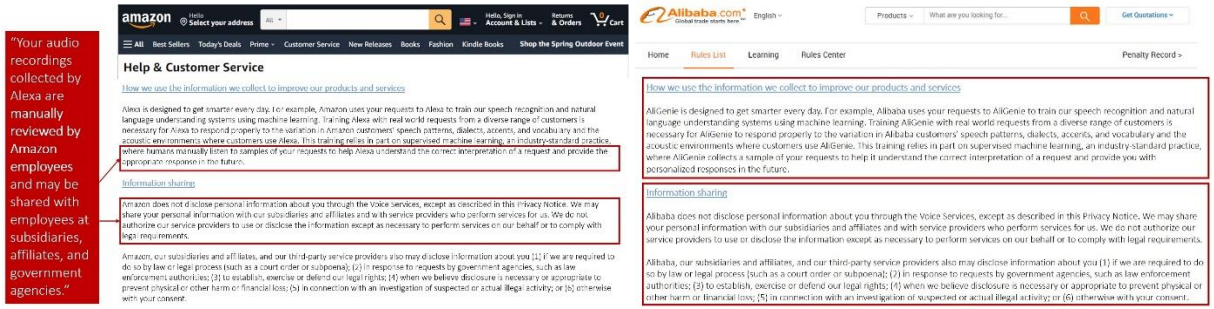
Ambivalence	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about interactions with **[Amazon/Alibaba]** (the company) . . ." 1. Possessed reactions toward **[Amazon/Alibaba]** that were mixed versus one-sided. 2. Felt conflicted in your reactions to **[Amazon/Alibaba]** 3. Experienced behavioral indecision **toward [Amazon/Alibaba]** 4. Felt tension in your thoughts and feelings toward **[Amazon/Alibaba]** 5. Felt ambivalent toward **[Amazon/Alibaba]** [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	"Indicate your agreement with the following statements:" 1. Possessed reactions towards the **seller** that were mixed versus one-sided. 2. Felt conflicted in your reactions to the **seller** 3. Experienced behavioral indecision 4. Felt tension in your thoughts and feelings towards the **seller** 5. Felt ambivalent towards the **seller**	Adapted from (Moody et al., 2014)
Response set item	**Prompt**: "If you are not a fish, then select "Agree" as the response to this question."	N/A	
Prior privacy experience	**Prompt**: "Read carefully and answer the following questions:" 1. How often have you personally experienced incidents whereby your personal information was used by some company without your authorization? 2. How often have you heard or read during the last year about the use and potential misuse of the information collected from virtual assistants (e.g., Alibaba AliGenie, Apple Siri, Amazon Alexa, Google Assistant, and the like)? 3. Same [Likert-type 7-point scale: 1 = Never; 7 = Always]	1. How often have you personally experienced incidents whereby your personal information was used by some company **or e-commerce web site** without your authorization? 2. How much have you heard or read during the last year about the use and potential misuse of the information **collected from the Internet?** 3. How often have you personally been the victim of what you felt was an improper invasion of privacy?	Adapted from Smith et al. (1996)
Awareness (of privacy practices)	**Prompt**: "Read carefully and indicate your agreement with the following statements:" 1. Same 2. Same 3. Same [Likert-type 7-point scale: 1 = Strongly disagree: 7 = Strongly agree]	1. Companies seeking information online should disclose the way the data are collected, processed, and used. 2. A good consumer online privacy policy should have a clear and conspicuous disclosure. 3. It is **very** important to me that I am aware and knowledgeable about how my personal information will be used.	Adapted from Malhotra et al. (2004)
Past positive experience	**Prompt**: "Overall, how you would rate the quality of Amazon Alexa?"	1. My past experience in **Amazon's auction marketplace** was positive	Adapted from Pavlou and

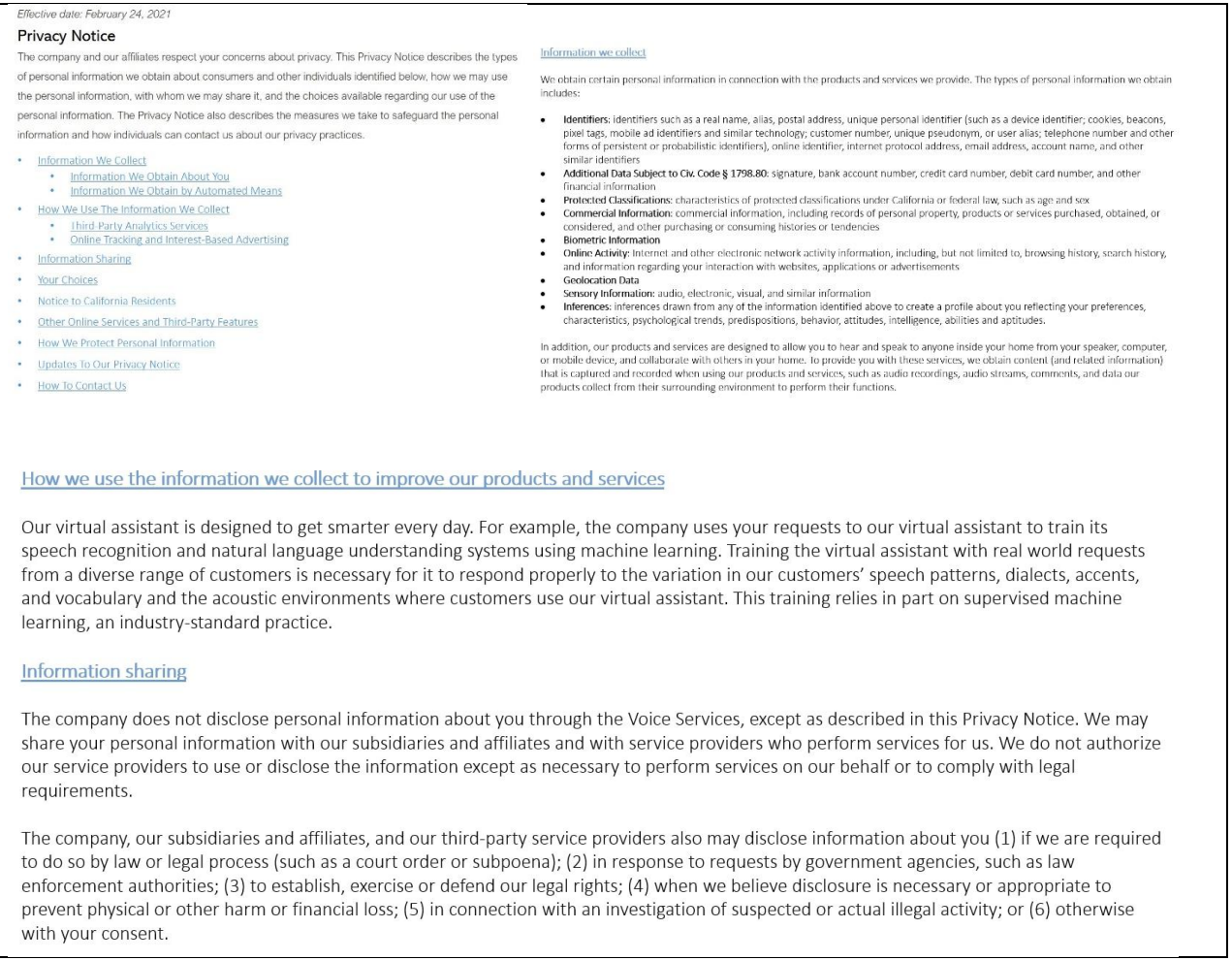
with Alexa	(Fill in the blank) 1. My past experience with **Amazon Alexa** was _____ 2. I received _____ service from **Amazon Alexa** in the past 3. **Amazon Alexa** did a _____ job in the past [Likert-type 5-point scale: 0 = No experience (N/A) 1 = Terrible: 5 = Excellent]	2. I received excellent service from sellers in **Amazon's auction marketplace** in the past 3. Sellers in Amazon's auction marketplace did a good job in the past	Gefen (2004)
Past positive experience with AliGenie	**Prompt**: "Overall, how you would rate the quality of Alibaba AliGenie?" (Fill in the blank) **1.** My past experience with **Alibaba AliGenie** was _____ **2.** I received _____ service from **Alibaba AliGenie** in the past 3. Alibaba AliGenie did a _____ job in the past [Likert-type 5-point scale: 0 = No experience (N/A) 1 = Terrible: 5 = Excellent]	1. My past experience in **Amazon's auction marketplace** was positive 2. I received excellent service from sellers in **Amazon's auction marketplace** in the past 3. Sellers in Amazon's auction marketplace did a good job in the past	Adapted from Pavlou and Gefen (2004)

1.9. Appendix 1C – Details on Experimental Treatments

Table C1. Description of Operationalized Variables and Levels

Construct	Operationalized variable	Manipulation description, levels, and stimulus
Default trust bias: innate, instinctual heuristic people develop during infancy that lessens conscious thought processing to evaluate the trustworthiness of others.	Product type	1) *Amazon Alexa (Echo)* – Alexa is Amazon's cloud-based voice service installed on hundreds of millions of smart devices from a variety of manufacturers including Amazon. 2) *Alibaba AliGenie (Tmall Genie)* – AliGenie is Alibaba's cloud-based voice service installed on Tmall devices with skills and applications that can be integrated into other hardware solutions.  
	Cloud storage provide (for voice data)	1) *Amazon web services* – Amazon is a U.S. based company providing cloud computing, hosting, and artificial intelligence solutions and has the most extensive global cloud infrastructure in the world. 2) *Alibaba cloud* – Alibaba is a Chinese-based company providing cloud computing, hosting, and artificial intelligence solutions to enterprises, organizations, and countries in more than 200 countries and regions.  

<table>
<tr><td rowspan="3">Salience bias: bias favoring salient over difficult, diffuse information, leading to suboptimal decision making.</td><td rowspan="3">Perceived human involvement in voice systems</td><td>Privacy policy excerpts with explicit reference to either human or machine processing of voice data:
How we use the information we collect to improve our products and services
[Alexa/AliGenie] is designed to get smarter every day. For example, [Alibaba/Amazon] uses your requests to [Alexa/AliGenie] to train our speech recognition and natural language understanding systems using machine learning. Training [Alexa/AliGenie] with real world requests from a diverse range of customers is necessary for [Alexa/AliGenie] to respond properly to the variation in [Alibaba/Amazon] customers' speech patterns, dialects, accents, and vocabulary and the acoustic environments where customers use [Alexa/AliGenie]. This training relies in part on supervised machine learning, an industry-standard practice . . .</td></tr>
<tr><td>1) Covert policy - Covert privacy policy does not mention explicit involvement of humans but references that voice data are used for product improvement and data analytics:
[(covert) where [Alexa/AliGenie] collects a sample of your requests to help it understand the correct interpretation of a request and provide you with personalized responses in the future].</td></tr>
<tr><td>2) Overt policy - Overt privacy policy mentions the explicit involvement of humans who will process participants' voice data for training data purposes:
[(overt) where humans manually listen to samples of your requests to help [Alexa/AliGenie] understand the correct interpretation of a request and provide the appropriate response in the future].</td></tr>
<tr><td>Level of effort: Level of cognitive effort or conscious awareness during decision making.</td><td>Level of effort (aid) in information processing</td><td>1) High level of aid – Callout boxes and written explanations of crucial privacy content expressing possible privacy implications for voice systems users.
2) Low level of aid – Callout boxes only as visual cues to highlight content of interest that may express possible privacy implications for voice systems users.
 
"Your audio recordings collected by Alexa are manually reviewed by Amazon employees and may be shared with employees at subsidiaries, affiliates, and government agencies."
High level of aid (left) vs. low level of aid (right)</td></tr>
</table>

Control condition:	*Omits* reference to: (1) a specific product type, (2) a specific cloud provider, (3) any specific agent (either human or machine) to process voice data for supervised training, (4) and any callout boxes or additional written aids	

Effective date: February 24, 2021

Privacy Notice

The company and our affiliates respect your concerns about privacy. This Privacy Notice describes the types of personal information we obtain about consumers and other individuals identified below, how we may use the personal information, with whom we may share it, and the choices available regarding our use of the personal information. The Privacy Notice also describes the measures we take to safeguard the personal information and how individuals can contact us about our privacy practices.

Information we collect

We obtain certain personal information in connection with the products and services we provide. The types of personal information we obtain includes:

- **Identifiers:** identifiers such as a real name, alias, postal address, unique personal identifier (such as a device identifier; cookies, beacons, pixel tags, mobile ad identifiers and similar technology; customer number, unique pseudonym, or user alias; telephone number and other forms of persistent or probabilistic identifiers), online identifier, internet protocol address, email address, account name, and other similar identifiers
- **Additional Data Subject to Civ. Code § 1798.80:** signature, bank account number, credit card number, debit card number, and other financial information
- **Protected Classifications:** characteristics of protected classifications under California or federal law, such as age and sex
- **Commercial Information:** commercial information, including records of personal property, products or services purchased, obtained, or considered, and other purchasing or consuming histories or tendencies
- **Biometric Information**
- **Online Activity:** Internet and other electronic network activity information, including, but not limited to, browsing history, search history, and information regarding your interaction with websites, applications or advertisements
- **Geolocation Data**
- **Sensory Information:** audio, electronic, visual, and similar information
- **Inferences:** inferences drawn from any of the information identified above to create a profile about you reflecting your preferences, characteristics, psychological trends, predispositions, behavior, attitudes, intelligence, abilities and aptitudes.

In addition, our products and services are designed to allow you to hear and speak to anyone inside your home from your speaker, computer, or mobile device, and collaborate with others in your home. To provide you with these services, we obtain content (and related information) that is captured and recorded when using our products and services, such as audio recordings, audio streams, comments, and data our products collect from their surrounding environment to perform their functions.

How we use the information we collect to improve our products and services

Our virtual assistant is designed to get smarter every day. For example, the company uses your requests to our virtual assistant to train its speech recognition and natural language understanding systems using machine learning. Training the virtual assistant with real world requests from a diverse range of customers is necessary for it to respond properly to the variation in our customers' speech patterns, dialects, accents, and vocabulary and the acoustic environments where customers use our virtual assistant. This training relies in part on supervised machine learning, an industry-standard practice.

Information sharing

The company does not disclose personal information about you through the Voice Services, except as described in this Privacy Notice. We may share your personal information with our subsidiaries and affiliates and with service providers who perform services for us. We do not authorize our service providers to use or disclose the information except as necessary to perform services on our behalf or to comply with legal requirements.

The company, our subsidiaries and affiliates, and our third-party service providers also may disclose information about you (1) if we are required to do so by law or legal process (such as a court order or subpoena); (2) in response to requests by government agencies, such as law enforcement authorities; (3) to establish, exercise or defend our legal rights; (4) when we believe disclosure is necessary or appropriate to prevent physical or other harm or financial loss; (5) in connection with an investigation of suspected or actual illegal activity; or (6) otherwise with your consent.

1.10. Appendix 1D – Additional Details on Results and Analyses

Table D1. Descriptive Statistics of Participants

Variable	**Category**	**Frequency (percentage)**
		n = 499
Gender	Male	250 (50.1%)
	Female	239 (47.9%)
	Prefer not to say	7 (1.4%)
	Nonbinary	3 (0.6%)
Age	18–24	48 (10%)
	25–29	70 (14%)
	30–34	108 (22%)
	35–39	80 (16%)
	40–49	104 (21%)
	50–65	78 (16%)
	Older than 65 years	11 (2%)
Annual income range	< $30,000	120 (24%)
	$30,001 – $75,000	207 (41.5%)
	$75,001 – $150,000	141 (28.3%)
	$150,001 – $300,000	27 (5.4%)
	$300,001 – $500,000	2 (0.4%)
	$500,001+	2 (0.4%)
Employment status	Employed part-time	65 (13%)
	Employed full-time	306 (61.3%)
	Not employed	42 (8.4%)
	Self-employed	54 (10.8%)
	Student	17 (3.4%)
	Retired	10 (2.0%)
	Other	5 (1.0%)
Educational level	Less than high school	0 (0%)
	High school graduate or equivalent	50 (10%)
	Some college but have not completed a degree	107 (21.4%)
	Associate degree or two-year equivalent	51 (10.2%)
	Bachelor’s degree	214 (42.9%)
	Master’s degree	71 (14.2%)
	Doctorate / Ph.D.	6 (1.2%)
Ethnicity	American Indian or Alaskan Native	2 (0.4%)
	Asian	46 (9.2%)
	Black or African American	47 (9.4%)
	Hispanic or Latino	36 (7.2%)
	Middle Eastern or North African	3 (0.6%)
	Native Hawaiian or Pacific Islander	0 (0%)
	White or Caucasian	358 (71.7%)
	Prefer not to say	7 (1.4%)
Familiarity with voice assistants	Yes	499 (100%)
Frequency of use of voice assistants	Less than once a month	20 (4%)
	1–3 times a week	120 (24%)
	Once a day	78 (15.6%)
	2–4 times a day	173 (34.7%)
	More than 4 times a day	108 (21.6%)
Mobile skills	Fair	15 (3%)
	Good	94 (18.8%)

	Very Good	194 (38.9%)
	Excellent	196 (39.3%)
Length of time of smart device ownership	Less than 12 months	23 (4.6%)
	13–24 months	90 (18%)
	25–36 months	88 (17.6%)
	37–48 months	51 (10.2%)
	49–60 months	36 (7.2%)
	61–72 months	28 (5.6%)
	More than 72 months	183 (36.7%)
Mobile operating system	iOS	249 (49.9%)
	Android	245 (49.1%)
	Windows	3 (0.6%)
	Other	2 (0.4%)
Updated mobile operating system in last 12 months	Yes	442 (88.6%)
	No	57 (11.4%)
Quality of experience with AliGenie (6-point scale: 1 No experience, 2 – 6 Terrible to Excellent)	Mean = 1.86	
	Median = 1.00	
	Std. Dev = 1.52	
Quality of experience with Alexa (6-point scale: 1 No experience, 2 – 6 Terrible to Excellent)	Mean = 4.62	
	Median = 5.00	
	Std. Dev = 1.12	

Table D2. Item Loadings

Identified factor	Item	Loading
Willingness to share voice data for processing	DV_1	0.881
	DV_2	0.928
	DV_3	0.861
Perceived human involvement	MCHum1_1	0.833
	MCHum1_2	0.757
	MCHum2_1	0.904
	MCHum2_2	0.849
	MCHum2_3	0.849
Benevolence / integrity	Benev_1	0.881
	Benev_2	0.747
	Benev_3	0.809
	Integ_1	0.89
	Integ_2	0.923
	Integ_3	0.83
	Integ_4	0.901
Competence	Compt_1	0.887
	Compt_2	0.825
	Compt_3	0.89
	Compt_4	0.786
Privacy concern	Intru_1	0.888
	Intru_2	0.895
	Intru_3	0.906
	SecUse_1	0.917
	SecUse_2	0.914
	SecUse_3	0.904
	Survei_1	0.668
	Survei_2	0.936
	Survei_3	0.896

Privacy risk	Risk_1	0.928
	Risk_2	0.897
	Risk_3	0.921
	Risk_4	0.886
Perceived benefits	Bene_1	0.876
	Bene_2	0.892
	Bene_3	0.745
Prior privacy experience	Prior_1	0.881
	Prior_2	0.617
	Prior_3	0.867
Privacy policy awareness	Aware_1	0.836
	Aware_2	0.875
	Aware_3	0.63
Marker variable	Mark_1	0.677
	Mark_2	0.838
	Mark_3	0.809

1.11. Appendix 1E – Mediation Testing

Traditionally, the Baron and Kenny (1986) and Sobel (1982) tests have been used to test for mediation. Because of the increased computing power available to researchers, other methods to test for mediation are becoming prevalent. Namely, the bootstrapping method has become the leading approach. First developed in behavioral research (Hayes, 2009; MacKinnon, 2008), it was recently introduced to the IS discipline (e.g., Vance et al., 2015). The advantages of this approach include the following: greater statistical power, direct measurement of "indirect effects," and no requirement to assume the data are normally distributed.

We performed the bootstrapping method by resampling (from the obtained sample) with replacement 5,000 times (Hayes, 2009). To estimate the indirect effect in each resample, we must obtain the product (ab) by multiplying the coefficients in paths a (i.e., independent variable → mediating variable) and b (i.e., mediating variable → dependent variable) (MacKinnon, 2008). Next, we must obtain the path coefficient from the independent variable to the dependent variable, which is the coefficient corresponding to c′. The percentile-based confidence interval ci% is calculated by sorting the values of ab and c′ in ascending order. We then used the formula k(.5 - ci/200) for the lower bound and the formula 1 + k(.5 + ci/200) for the upper bound to calculate the

ordinal positions of ab and c′ corresponding to the bounds of our interval, where k is the number of resamples (Hayes, 2009). Assuming a 95% confidence interval, our calculated ordinal ranges were 125 and 4,876 for the lower and upper bounds, respectively.

Finally, to determine if an indirect effect existed, we observed the confidence interval ab. If the upper and lower bounds do not include zero, then we can conclude with a confidence of ci% that the indirect effect existed and that it was not zero (MacKinnon, 2008). Furthermore, we can determine whether the mediation is full or partial by examining the confidence interval for c′. If the confidence interval for ab does not include zero, but the confidence interval for c′ does include zero, then the effect is fully mediated. Conversely, if the confidence intervals for both ab and c′ do not include zero, then the effect is partially mediated.

1.12. Appendix 1F – Human Involvement Qualitative Comments

Table F1. Open-ended Responses on Human Involvement Concerns

#	**Participant comment**
1	I don't like the idea of humans listening to my conversations.
2	I did not realize that humans were involved in data collection
3	I do not like that actual humans are listening to my commands
4	I don't really mind if they have my voice as long as they are not listening and storing personal data or conversations, which I can neither prove nor disprove they do
5	I don't like that they use humans to decide what the outcome is
6	analyze this information to farm out to other 'subsidiaries' without me even knowing about where it's going or who's listening
7	Companies will do unethical things to make more money and there is the potential for a lot of money to be made listening to records and selling those records to other people or companies.
8	I do not like that my information would be shared with other humans.
9	The idea of having humans listening to my voice recordings, even in the context of making the service better, is really unsettling. I realize on some level that this is probably standard across all personal assistant devices, but actually reading the privacy policy and thinking back to conversations I've had around my personal assistant devices makes me concerned for what exactly is out there of my personal information.
10	I have absolutely no interest in letting humans of any kind listen to my data, my private and personal moments at home, or even just when I'm interacting with this thing. That's horrifying. It feels like eavesdropping.
11	I do not want to run the risk of others hearing and reviewing my interactions with the device, I feel it is an invasion of my privacy.
12	would not want to share my voice data with Alibaba because I would be embarrassed by the thought of someone hearing me speak to the machine even if my identity were not known. I would feel like I was being listened in on and would not use the device as much or as freely.
13	This are third party people who are listening. Its private information and I don't feel comfortable with third party people listening to it. This something I am not happy with. Know a days people can use this against you.
14	It really makes me weirded out considering that people could then frame someone for any crime and use their

#	**Participant comment**
	OWN voice as a confession.
15	It feels a bit weird for people to have voice recordings of me, even if it is a company with reasonable intentions.
16	I also really dislike the idea of real people manually going through the voice data.
17	The fact that I know humans will be listening to my voice and that it can be shared with a third party is a bit more disconcerting.
18	nor should they have humans review it for AI training
19	correct way to think about this is to assume that anything you say to your digital assistant might very well be heard by someone else in the future.
20	It seems sketchy to me that the voice data might be shared with humans. It seems like an invasion of privacy.
21	but I would not like to them check my input
22	It's very creepy knowing my voice is recorded and listened by someone else that I do not know. It's one thing speaking to someone in person but not having your voice recorded, saved and listened by another.
23	Last, what if I'm talking about confidential things that I do not want others to hear and Alexa recorded it and again someone else is listening to it.
24	My voice data is unlikely to be abused based on the privacy policy, which indicates that Amazon primarily uses it to improve Alexa, and no human actor is used to listen to my conversations.
25	I do not like that my voice will be listened to by employees of Alibaba and would not agree to this invasion of privacy.
26	On Google Nest we can stop storage activity if we want to. I have never read Google "human employees" listen to my voice to make Google Nest more efficient and would be shocked to read this.
27	I don't like the idea of Alexia recording my conversations and sending them to be stored at Amazon, to be listened to by who knows.
28	Personally, I would prefer not to share my voice data with Alibaba because I am concerned about my privacy. I was put off by the fact that voice recordings are listened to and manually interpreted by human agents at Alibaba.
29	I would not want to share my voice data with Alibaba because I take issue with the way that it would be used and shared. I would not want a human to listen to my data,
30	This assistant is in my home hearing all kinds of conversations and I would not like to share all of this with others. It seems way too open to others to me.
31	I wouldn't be prone to exercise these feature as stated before others will have access to my voice recognition.
32	I don't need people listening to my recorded voice to supposedly help the Alibaba Genie interpret what I am saying. I also don't want my personal information of address, bank account number, or credit card given to them either.
33	I would be worried they would use my voice for other purposes and others would hear my voice.
34	I find data analysis to be very privacy invasive, I wouldn't feel comfortable with that much information about me being available to unknown people
35	I don't want anyone to be able to access my personal information,
36	I don't want them to sit and listen to my voice data because it's none of their business.

Chapter 2: Beyond Privacy Concerns: The Conceptualization and Measurement of Privacy Interest

2.1. Introduction

Privacy protection in today's automation age has imperceptibly diminished over time. Consumer data are ubiquitously collected and commoditized in online data markets and company data stores (Spiekermann et al., 2015), often for consumer profiling (Al-Natour et al., 2020; Neumann et al., 2019). The economic value of consumer data has fueled a digital advertising industry expected to reach nearly $2 trillion globally in the next decade (Research and Markets, 2022). Data such as biometrics (Breward et al., 2017; Du et al., 2020), connected vehicles (Cichy et al., 2021), location tracking (Crossler & Bélanger, 2019), social relationships (Choi et al., 2018), and sexual data (Maris et al., 2020) are leveraged by companies to improve their predictions of consumers' behavioral outcomes (Zuboff, 2015, 2019)—a practice known to heighten consumers' concerns for information privacy (Pew Research Center, 2019). Although "privacy concerns are at an all-time high" (Zhang et al., 2022, p. 492), consumers' behaviors do not match their elevated privacy concerns, a phenomenon referred to as the privacy paradox (Acquisti & Gross, 2006; Norberg et al., 2007). As such, government regulations on information privacy protection in the US have been slow to pass given that consumers continue to use the same technologies that elevate their privacy concerns (Solove, 2021; Westin, 2000). As a result, protection of consumers' privacy is often the responsibility of the individual (Acquisti et al., 2022).

The concept of privacy is multidimensional and situational (Laufer & Wolfe, 1977). Laufer and Wolfe (1977) described the dimensions of privacy and the "elements of situations" (p. 25) that influenced people's perceptions of privacy and invasion, explaining the types of situations that affected people's privacy perceptions, why these situations were perceived as invasive, and why certain people were aware of these privacy-related experiences. Information systems (IS)

researchers updated the concept of privacy in the digital age to *information privacy*—broadly defined as the ability to control one's personal information (Bélanger & Crossler, 2011). As technologies became more powerful and sophisticated, the conceptual domain of information privacy further evolved (Smith et al., 2011). The dimensionality of privacy changes because the perception of privacy changes over time (Bansal & Nah, 2022; Hong & Thong, 2013; Smith et al., 1996; Xu et al., 2012a). During the rise of the Internet, Smith et al. (1996) created the concern for information privacy (CFIP) scale that included the dimensions of *collection*, *errors*, *unauthorized secondary use*, and *improper access*. Malhotra et al. (2004) introduced the internet users' information privacy concerns (IUIPC) scale, positing that a new set of privacy dimensions—*awareness of privacy practices*, *control*, and *collection*—were at the center of consumers' privacy concerns. As mobile internet and social media became more ubiquitous, Xu et al. (2012a) proposed the mobile users' concerns for information privacy (MUIPC) scale, arguing that users' perceptions of *surveillance* and *intrusion* and concerns over *secondary use of information* elicited greater privacy concerns. Finally, Hong and Thong (2013) proposed the internet privacy concerns (IPC) scale by reconceptualizing the dimensions posited by Smith et al. (1996) and Malhotra et al. (2004) as a third-order factor comprising two second-order factors (interaction management, information management) and six first-order factors (awareness, errors, control, collection, secondary usage, improper access). As new internet technologies became enmeshed in consumers' daily lives, prompting new privacy needs, researchers expanded their conceptualizations of privacy concern.

The development of privacy concern scales was based on the shared understanding that privacy is a situational phenomenon evolving over time. This is contrary to a normative perspective of privacy decision-making. Through the normative privacy lens, consumers would be viewed as rational actors who perform privacy behaviors in concert with their privacy beliefs, attitudes, and

intentions (Culnan & Armstrong, 1999; Dinev & Hart, 2006) and are unaffected by situational factors. That is, they make optimal privacy calculus decisions (i.e., a rational trade-off decision on the risks and benefits of information disclosure) in every privacy context. By contrast, the behavioral privacy perspective explains that consumers would be bounded by cognitive biases and make privacy decisions based on situational cues (Acquisti, 2004; Acquisti et al., 2015; Acquisti et al., 2016) and thus cannot make actual decisions that reconcile with their attitudes and beliefs in every privacy context (Acquisti & Gross, 2006; Adjerid et al., 2018b; Norberg et al., 2007). An issue then arises when researchers adopt a normative privacy perspective and assume consumers' stated privacy concerns are robust against situational influences and reliably predict actual behaviors, but then apply a privacy concern scale with dimensions that are "neither absolute nor static" and that represent the construct at the time of the scale's development (Smith et al., 1996, p. 190). When a mismatch occurs between consumers' responses to a privacy concern scale and their actual behaviors, this is called the privacy paradox (Acquisti et al., 2020). Thus, developing a privacy measure that is more robust to situational influences would be helpful to practice and research in advancing future conceptualizations of information privacy.

The research aim of this study is to reinvestigate the "theoretical and operational assumptions" of privacy concern and offer a reconceptualized construct called *privacy interest* "in light of emerging technology, practice, and research" (Stewart & Segars, 2002, p. 37). Drawing on the cognitive model of empowerment (Thomas & Velthouse, 1990), we posit that privacy interest focuses on a consumer's dispositional interest toward privacy, which is in contrast to the situationally-oriented privacy concern construct that measures consumers' reactions to privacy events or companies' data practices. Originally theorized in the extended privacy calculus, interest was included as a predictor on the confidence and enticement beliefs (benefits) side of the equation

to predict consumers' behavioral intentions (Dinev & Hart, 2006). Dinev and Hart (2006) conceptualized interest as *personal internet interest* (i.e., intrinsic motivation to use the Internet) and directly contrasted it with privacy concern to understand which was greater. Although Dinev and Hart (2006, p. 74) stated "personal interest enriches the privacy calculus model and should be included in future models," interest was rarely used in studies operationalizing the privacy calculus (see Appendix 2A for literature review). Therefore, we reposition interest on the risk belief side of the privacy calculus equation and theorize it as an interest in information privacy protection developed in stages from a situational interest to a dispositional interest in privacy, and we operationalize interest through the dimensions of *awareness*, *meaningfulness*, *impact*, and *competence*.

The cognitive model of empowerment was originally conceptualized in a workplace environment (Spreitzer, 1995; Thomas & Velthouse, 1990) and was later reconceptualized to student interest (Frymier et al., 1996; Weber & Patterson, 2000). The dimensions that readily apply to a workplace and academic environment also apply to a privacy context. The dimensions we include in our recontextualized privacy interest are as follows. First, we replace the original dimension of choice with *awareness*. Often, consumers are not given a choice to preserve their privacy if they wish to use essential digital technologies (Alashoor et al., 2022). In a privacy context, we assert that having awareness of a threat to one's privacy or of one's lack of choice in a privacy situation serves as a situational trigger to engage in privacy protection behaviors. Second, *meaningfulness* reveals how much a person values privacy. This dimension is similar to disposition to value privacy, an antecedent to privacy concern (Dinev et al., 2015; Smith et al., 2011), and is an important consideration in privacy decision-making when consumers become saliently aware their privacy is under threat. Third, *competence* relates to whether consumers have the self-efficacy

and confidence to successfully protect their privacy. Consumers who feel that protecting their privacy is a meaningful endeavor will develop the competence and eventually the confidence to successfully safeguard their personal information. Fourth, *impact* is an important consideration. Consumers may not feel their individual-level actions make a difference in preventing companies from collecting and using their data. Each dimension has an additive effect on how a consumer assesses a situation (Thomas & Velthouse, 1990), and thus those who value their privacy and who have the competence to safeguard it will assume their actions contribute to their privacy protection. Considering that the dimensionality of privacy interest differs from those in previous privacy concern scales and that the assessment of people's dispositional attitudes toward privacy differs from their situational reactions to privacy events, we maintain that an updated conceptualization of privacy is necessary as societies become increasingly automated.

We follow MacKenzie et al. (2011) and develop a scale for privacy interest to identify largely stable privacy dimensions that are not situationally dependent (i.e., a person's response to an event). We define *privacy interest* as the general feeling toward reengaging protective behaviors that increase one's information privacy. Our study posits that users who have higher levels of privacy interest, manifested in the four dimensions, will engage in proactive privacy protection behaviors, specifically mobilization efforts that can motivate political action to address the harms created in a surveillance economy (Mulligan et al., 2020).

Our study makes several contributions to the information privacy discourse. First, we refocus on the extended privacy calculus and expand the conceptualization of interest to an interest in information privacy and protection. We establish the dispositional link to privacy decision-making by clarifying the role privacy interest has in the privacy calculus. Second, our conceptualization of privacy interest is a more robust measure in situational contexts, especially contexts that artificially

raise privacy concern perceptions and lead to a misconstruing between attitudes and behaviors (Preibusch, 2013). For instance, established scales may trigger concerns related to data misuse when participants are asked to think through the dimensions of privacy concerns (i.e., control, access, error, use, collection, awareness) when they otherwise would not have considered them (Cichy et al., 2021). In addition, the dimensions of privacy interest are reconceptualized in a privacy environment that reflects today's automation age regarding (un)awareness of privacy threats, meaningfulness of one's privacy, competence to perform protective behaviors, and impact of one's actions. Third, the formation of the privacy interest construct provides another theoretical lens from which to view and operationalize the privacy calculus theory. Finally, our study contributes to the information privacy discourse through the development of a valid and reliable measurement scale to facilitate future empirical research on privacy.

2.2.Theoretical Background

2.2.1. Cognitive Model of Empowerment

Privacy protection is inherent in human nature (Acquisti et al., 2022; Acquisti et al., 2020; Altman, 1975, 1977). Humans are intrinsically motivated to protect their privacy, so much so that protection behaviors go unnoticed because they occur so pervasively (Acquisti et al., 2020). For instance, humans perform privacy protection behaviors with little conscious awareness, ranging from adjusting their voice levels in a group setting, covering documents they read, taking personal calls in a separate room, to closing blinds in their homes at night (Acquisti et al., 2022; Acquisti et al., 2020; Altman, 1977). In an online context, humans adjust their privacy settings on social media accounts, toggle cameras and microphones on/off during conference calls, and respond selectively to intended parties in their email correspondence in an attempt to regulate boundaries with others (Acquisti et al., 2020).

Maintaining one's privacy can be understood as a form of intrinsic task motivation through the

cognitive model of empowerment (Thomas & Velthouse, 1990). *Empowerment* refers to the changes in task assessments (i.e., cognitive elements) that affect one's intrinsic task motivation. *Intrinsic task motivation* refers to the satisfaction or positive experiences one derives from a task and is manifested in four cognitive elements: choice, meaningfulness, impact, and competence. Formally, a person performs *task assessments* through the four cognitive elements to determine whether a task (i.e., any activity performed with a purpose in mind) elicits an intrinsic task motivation to continue performing the task. The cumulative learning from past task assessments is referred to as *global assessments*, which are formed over time and can influence future task assessments. In a workplace setting, intrinsic task motivation reflects "an individual's orientation to his or her work role" (Spreitzer, 1995, p. 1443) and serves as a barometer of the level of empowerment employees feel in performing their work roles. **Table 1** summarizes the key theoretical dimensions and definitions of the cognitive model of empowerment.

Table 1. Dimensions of Cognitive Empowerment

Dimension	**Definition (Thomas & Velthouse, 1990)**	**Application in an information privacy context**
Task	The performance of an activit(ies) with a specific purpose in mind.	The performance of behaviors that protect one's information privacy.
Meaningfulness	The value or intrinsic caring a person has toward the task (purpose).	The value of performing privacy behaviors in relation to one's beliefs.
Global meaningfulness	The general (aggregate) level of caring, commitment, or psychological investment a person has in the task (purpose).	
Impact	Degree to which a person can "make a difference" in the outcome of the task (purpose).	Degree to which one's behavior makes a difference in protecting their information privacy.
Global impact	The expectancy of having an impact on the task, which is subject to change over time to include new experiences.	
Competence	Degree to which a person can skillfully perform the task (also referred to as self-efficacy; Bandura, 1977).	Degree to which one feels capable to perform the necessary activities to protect their information privacy.
Global competence	The general sense a person has in their ability to perform similar tasks reasonably well in new environments or circumstances (also referred to as self-confidence; Wells & Marwell, 1976).	
Choice	The belief a person is causally responsible for one's behavior (as origin or pawn; DeCharms, 1968).	Degree to which one feels' responsible for their privacy actions.
Global choice	The general tendency to view oneself as origin or pawn, which is subject to change over time based on life experiences.	

2.2.2. Empowerment Operationalized as Interest

Outside of the workplace, empowerment has been operationalized as interest (Weber et al., 2005; Weber & Patterson, 2000). *Interest* is considered "the psychological state of engaging or the predisposition to reengage with particular classes of objects, events, or ideas over time" (Hidi & Renninger, 2006, p. 112). In the education sciences, the learner empowerment scale (LES) was created that relied on the cognitive model of empowerment (Thomas & Velthouse, 1990) to assess learner's motivation to study (Frymier et al., 1996). Specifically, student motivation to learn in a classroom environment is attributed to their *state* and *trait* motivations (Brophy, 2004; Frymier et al., 1996). From an empowerment perspective, state motivation is derived from task assessments and refers to a student's desire to learn specific content at a specific point in time; trait motivation is derived from global assessments and refers to a student's inherent drive to perform the task of learning because the task itself is meaningful (Frymier et al., 1996; Thomas & Velthouse, 1990). Researchers examined the conceptual similarities between the multidimensional LES and a unidimensional Perceived Interest Questionnaire (Schraw et al., 1995) and found the LES to be a more effective, more valid and reliable measurement of student interest (Weber et al., 2005; Weber & Patterson, 2000).

Notably, interest develops from a situational interest to an individual (dispositional) interest and occurs in four stages (Hidi & Renninger, 2006; Renninger & Hidi, 2019). *Situational interest* refers to a heightened state of focused attention and affective reaction caused by an environmental stimuli (trigger) that may dissipate over time, whereas an *individual interest* refers to the predisposition to reengage the stimuli that originally initiated the heightened state of attention and affective reaction (Hidi & Renninger, 2006). In stage 1, *triggered situational interest*, interest is typically triggered by the awareness of an environmental feature that causes a short-term change in a person's cognitive and affective processing. Stage 2, *maintained situational interest*, involves

an extended cognitive or affective processing of the trigger that occurs over time or in subsequent episodes. The meaningfulness of the task holds and sustains the situational interest in this stage (Harackiewicz et al., 2000; Mitchell, 1993). Stage 3, *emerging individual interest*, is based on previous task assessments and is the start of an emerging predisposition to reengage with a task over time. In this stage, a person develops stored value and knowledge (competence) of the task. Stage 4, *well-developed individual interest*, is a global assessment where a person performs tasks seemingly effortlessly and spends more time on a task, producing "more types and deeper levels of strategies for work with tasks" (Hidi & Renninger, 2006, p. 115). In summary, people require a situational trigger to spark interest in a task, and through the stages of development, this interest becomes a predisposition to reengage the task over time and in subsequent episodes.

2.2.3. Existing Conceptualizations of the Privacy Calculus

The conceptualization of task and global assessments and interest development serves as the foundation of our discussion on privacy calculus. Culnan and Armstrong (1999, p. 106) defined privacy calculus as an assessment customers make when they disclose personal information to retailers that their data "will subsequently be used fairly and they will not suffer negative consequences" (Laufer & Wolfe, 1977; Milne & Gordon, 1993; Stone & Stone, 1990). The context involved customers' provision of transaction data and other personal information for use in retailers' targeted marketing campaigns in exchange for a social or economic benefit. Dinev and Hart (2006) extended the privacy calculus with respect to the Internet, focusing on an ecommerce context involving vendors and online transactions. In the age of automation, interactions with companies have evolved and are no longer confined to an online transactions environment. Consumer-company interactions can occur at a person's workplace or home, inside their car or bedroom, on a phone or wearable device—nearly anywhere with an internet connection. However, people's calculus to weigh the benefits versus costs of using emerging technologies will differ

depending on how conversant one is in anticipating the downstream privacy cost of technology use—which is arguably impossible to estimate (Acquisti et al., 2020). As updates to previous privacy decision-making models, the Antecedents → Privacy Concerns → Outcomes (APCO) model (Smith et al., 2011) and later *enhanced* APCO (Dinev et al., 2015) were developed as macromodels to describe the entire privacy decision-making context, which included the privacy calculus and principles from behavioral economics and psychology. In the APCO macromodels, interest is omitted and privacy concern is the primary decision criterion bridging antecedents and outcomes.

Few studies have examined the role of interest a person has in either technology or in privacy when they perform a privacy calculus. **Table A1** (Appendix 2A) summarizes the research published in the Senior Scholars' List of Premiere Journals that cited the extended privacy calculus (Dinev & Hart, 2006). The studies that investigated consumers' interest pertained to an interest in information technology (Jia et al., 2022), in social commerce (Shen et al., 2019), the screening out of low-interest users through an ex ante registration request (Huang et al., 2021), and the covariation of interests in which interaction partners' interests corresponded (Al-Natour et al., 2021). A few studies outside the Basket of 11 have also examined digital interest (James et al., 2013) and personal app interest (James et al., 2021). None of the studies, however, examined whether having an interest in privacy would make consumers more resolute to protect their privacy. Research has examined interest in the "confidence and enticement beliefs" (benefits) side of the calculus but has left a gap in examining consumers' interest from the "risk beliefs" (risk) side of the calculus. For a person to perform a more effective calculus, a person's interest in learning the myriad downstream risks associated with emerging technologies would require a development of interest from situational triggers (i.e., privacy events) to a dispositional interest,

where competence in assessing risk and risk mitigation situations is necessary to perform a full calculus.

Risk beliefs are multidimensional and can range from physical to decisional (Karwatzki et al., 2022). The downstream consequences consumers can consider are similarly aplenty (Karwatzki et al., 2017b). To assess risk events, consumers must understand the impact of the event and the probability of its occurrence (Karwatzki et al., 2017b; Karwatzki et al., 2022). While researchers have admonished that expanding our understanding of privacy risks is crucial to advancing the privacy discourse (Dinev & Hart, 2006; Dinev et al., 2013), this admonishment did not include deepening our understanding of consumers' interest levels to expand their knowledge of information privacy, privacy events, and privacy protection. Our review shows that only a few studies have explored interest as a predictor in the privacy calculus to examine the benefits and conveniences afforded by the Internet and other technologies, and none have examined an interest in privacy. In an age where technologies are deployed without people's knowing (e.g., smart cities, surveillance technologies), privacy research could benefit from an increased focus on studying people's interest in privacy from the risk-beliefs side of the privacy calculus equation.

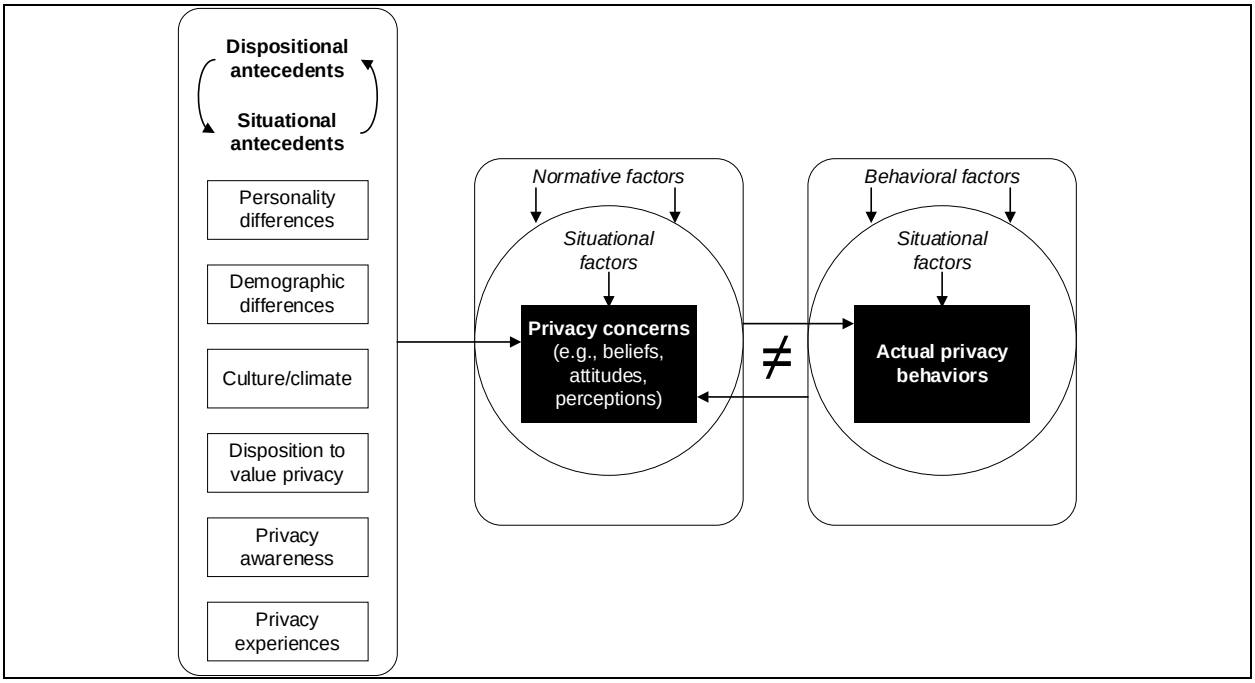
2.2.4. Existing Interpretations of the Privacy Paradox

Research has discussed the roles of normative and behavioral factors in privacy decision-making (e.g., Adjerid et al., 2018b; Bélanger & James, 2020; Laufer & Wolfe, 1977; Smith et al., 2011; Tsai et al., 2011). Normative factors influence the objective, rational privacy calculus of benefits and costs in privacy decision-making (Adjerid et al., 2018b). Examples of normative factors include privacy regulations, company privacy policies, information sensitivity, and privacy notices and seals. The normative perspective expects complete rationality. Behavioral factors do not influence the objective, rational privacy calculus of benefits and costs in privacy decision making but can affect actual behaviors (Adjerid et al., 2018b). Examples of behavioral factors include

choice defaults, cognitive resources, motivation, choice architecture, and time constraints. The behavioral perspective anticipates bounded rationality. Moreover, situational factors can influence the objective, rational privacy calculus of benefits and costs in privacy decision making *and* actual behaviors (Dinev et al., 2015). Examples of situational factors include biases, heuristic processing, misattribution, and cues and signals. As shown in **Figure 1**, a different set of normative, behavioral, and situational factors can simultaneously influence concerns and behaviors. When concerns do not accurately predict behaviors, this mismatch is considered a paradox. We believe that a major impediment to research on the privacy paradox is the absence of a privacy construct and measure that remain largely stable in a variety of privacy situations.

Our belief is that an expanded conceptualization on information privacy in the age of automation requires a perspective on privacy interest that is oriented toward identifying dispositional attributes of consumers who protect their privacy. By contrast, in the information age the privacy concern perspective focused on the environmental factors exerted on consumers that elicited their concerns. A perspective on interest is inherent in the privacy calculus but not on the privacy side of the equation. Thus, a focus on interest in privacy rather than on interest to use technologies provides an alternative lens through which to communicate how privacy risk events can affect consumers' dispositional privacy attitudes, which can lead to fewer paradox occurrences, help inform effective strategies to increase privacy protection competence, and contribute to the development of comprehensive privacy regulations focused on educational interventions. A focus on privacy interest deepens our understanding of information privacy and broadens our discussions on privacy in light of autonomous technologies.

Figure 1. Viewing the Privacy Paradox Through a Situational ⇆ Dispositional and Normative ⇆ Behavioral lens



Note. Privacy-related situational antecedents (e.g., privacy experiences) trigger a consumer's disposition toward privacy, which later becomes an engrained disposition over time (e.g., disposition to value privacy).

2.3. Conceptualization of Privacy Interest

2.3.1. Stages of Interest Development

According to Hidi and Renninger (2006), interest development is an evolutionary process that requires an environmental feature to elicit a situational interest that over time manifests dispositionally. In a privacy context, a triggered situational interest occurs when consumers experience a privacy event, such as receiving a notice from a company informing their data have been compromised in a recent breach, making them saliently aware their privacy is endangered. For privacy fundamentalists, they will enter the second stage of privacy interest development—maintained situational interest—given they ascribe high value to privacy (Westin, 2000), and meaningfulness of privacy holds and sustains situational interest in this stage (Harackiewicz et al., 2000; Mitchell, 1993). For the privacy unconcerned who place no to low value on privacy (Westin, 2000), their triggered situational interest would dissipate then subsequently conclude their interest development. The third stage of privacy interest development marks the beginning of an emerging

dispositional interest in privacy and its protection, characterized by a consumer who develops competence in privacy protection behaviors, values privacy, and is aware of the situations that pose risks to privacy. In the fourth stage—well-developed dispositional interest—consumers practice privacy-protection strategies effortlessly as part of their routine behaviors. They continually reengage privacy-protection behaviors because they find their actions make a difference in their privacy protection.

2.3.2. Development Process from Situational to Dispositional Privacy Interest

In an ambient computing environment, smart technologies are designed to be hidden (Olwal & Dementyev, 2022), and in privacy "what is hidden from us either individually or collectively can be potentially harmful" (Laufer & Wolfe, 1977, p. 23). Consumers may not be consciously aware of situations that endanger their privacy (Bélanger & Crossler, 2011; Laufer & Wolfe, 1977). As Acquisti et al. (2022) explain, consumers do not see companies and governments listening to their conversations or peering into their screens when browsing online because their physical space may signal strong territorial privacy—a personal space carved out for oneself (Porteous, 1976). However, when smart technologies are added to the space, the space transforms into a virtual, privacy-sensitive environment subject to data collection and surveillance. Therefore, before actively performing a privacy calculus, consumers must be aware that a possible privacy risk event exists, which serves as a situational trigger to prepare a possible privacy-protective behavior. Importantly, awareness can catalyze one's thinking toward privacy, which activates their disposition to value their privacy (i.e., the meaningfulness of their privacy). Taken together, situational privacy interest begins with awareness that one's privacy is possibly endangered, which then triggers one's thoughts of how meaningful their privacy is in the event of an intrusion.

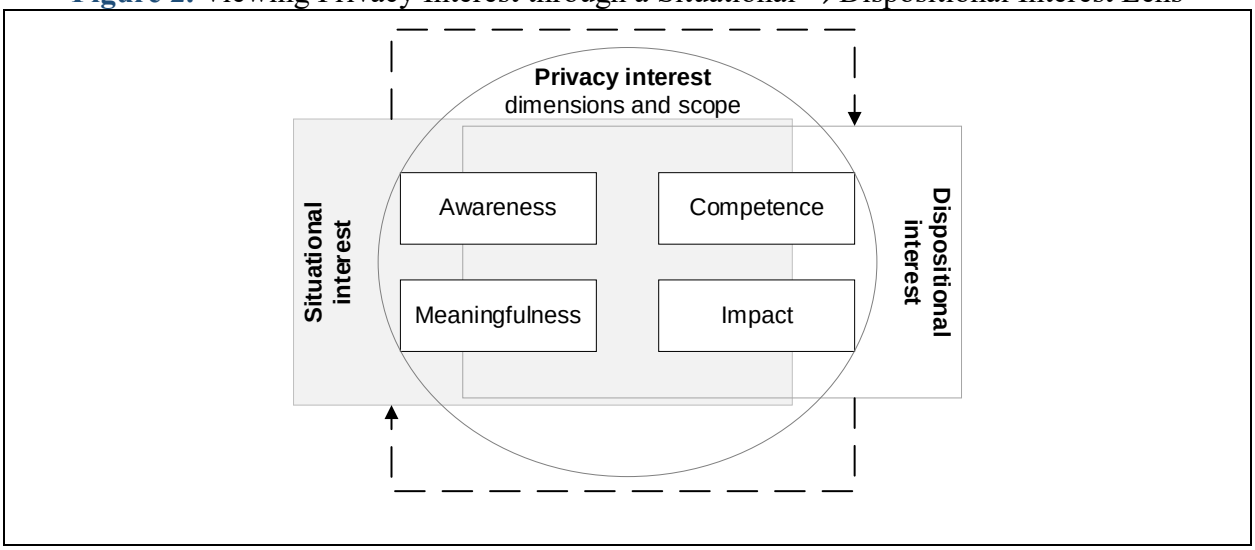
The APCO macromodel describes six antecedents to privacy concerns: (1) personality differences, (2) demographic differences, (3) culture, (4) disposition to value privacy, (5) privacy

awareness, and (6) privacy experiences (Dinev et al., 2015; Smith et al., 2011). Some antecedents are dispositional in nature (e.g., personality, demographics), and others are situationally specific to privacy (e.g., privacy awareness, privacy experiences), much like the dimensions of privacy interest. Hidi and Renninger (2006) explain that the characteristics of each stage of interest development function as mediators to the subsequent stage, deepening levels of interest sequentially. As such, when situational privacy interest becomes dispositional, consumers are likely to become resourceful in their protection strategies when in conditions fraught with privacy risks. In addition, consumers with a dispositional privacy interest will become anticipatory of privacy risks when they engage with or within environments with smart technologies, simultaneously considering the context and the appropriate strategies to mitigate privacy threats and risks. Finally, dispositional privacy interest is evident when the consumer perseveres to enact privacy protection behaviors, even when doing so is inconvenient or time-consuming. Although the APCO antecedents and privacy interest share common dimensions (i.e., awareness, meaningfulness), the cogency of the developmental process described in the stages of interest development provides support for a comprehensive conceptualization and measurement of dispositional privacy attitudes.

Drawing on the cognitive model of empowerment and the stages of interest development, we conceptualize privacy interest as a general feeling to reengage behaviors that protect information privacy manifested in a set of four cognitions that vary along the situational–dispositional continuum with (1) awareness and (2) meaningfulness situationally oriented and (3) competence and (4) impact dispositionally oriented. Formally, construct scope refers to "the set of things that possess the property represented by the construct" (Weber, 2021, p. 1648). The awareness and meaningfulness dimensions are within the scope of situational interest, whereas competence and

impact are within the scope of dispositional interest, and the scopes of both interests overlap and influence one another. **Figure 2** shows the dimensions and scopes of privacy interest and how they relate to one another.

Figure 2. Viewing Privacy Interest through a Situational ⇆ Dispositional Interest Lens



The four dimensions provide a comprehensive coverage of the scope of the privacy interest construct in the context of ambient computing (Olwal & Dementyev, 2022) and surveillance capitalization (Zuboff, 2015, 2019). Privacy interest includes the well-validated awareness dimension, common in Malhotra et al. (2004) and Hong and Thong (2013), and adds three new dimensions to the privacy calculus discourse (i.e., meaningfulness, competence, and impact). An interest in privacy is important to understand the extent to which a consumer is willing to develop stored knowledge of privacy protection strategies and to perform a privacy calculus more effectively, especially when privacy experiences become less salient and consumers become less aware of privacy risk events (Laufer & Wolfe, 1977).

2.3.3. Dimensions of Privacy Interest

In context-aware computing environments, sensor-embedded devices can perceive their environments to capture audio-visual data emitted from people and objects (Schuetz & Venkatesh,

2020). Personally identifiable information such as weight, height, appearance, voice, and the like are collected often without a consumer's awareness or permission (Wakefield, 2013). Awareness of a privacy situation is necessary for consumers to consciously choose to manage their privacy, and it is often in hindsight that they become aware of a privacy invasion event (Laufer & Wolfe, 1977). Privacy *awareness* refers to the extent to which a consumer is apprised of organizations' information privacy practices (Malhotra et al., 2004). Whether consumers choose to perform privacy-protective behaviors largely depends on their awareness of privacy issues in their surrounding environment (Belanger & Crossler, 2019).

Meaningfulness concerns the value of performing privacy behaviors with respect to one's beliefs and standards. Akin to disposition to value privacy, meaningfulness is the level of intrinsic caring consumers have toward the protection of their privacy. Thomas and Velthouse (1990) explained that higher levels of meaningfulness result in higher levels of involvement, commitment, and concerted effort toward a task, which in this case is privacy protection. By contrast, lower levels of meaningfulness result in apathy toward privacy protection, the boundary management of one's space, and the control of personal information flow (Xu et al., 2011a). As a disposition, global meaningfulness is the extent to which a consumer would invest in privacy protection behaviors. Thus, consumers with low levels of global meaningfulness will find the task of privacy protection a noncritical endeavor in the scope of their daily activities, whereas consumers with high levels of global meaningfulness will practice privacy protection behaviors routinely throughout their day.

Competence refers to the degree to which one feels capable of performing a task skillfully. In a privacy context, the task pertains to protection of information privacy and relates to a level of self-efficacy or personal mastery (Bandura, 1977, 1986) a person has to perform protective privacy

behaviors. Self-efficacy beliefs are "people's judgments of their capabilities to organize and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391). People with low self-efficacy beliefs avoid tasks that require high competence, thereby precluding them from building the necessary skills to perform the task in the future (Bandura, 1977). Conversely, people with high self-efficacy beliefs tend to exert high levels of effort to perform a task and will persevere through difficulties they encounter.

Impact, defined as the degree to which one's behavior makes a difference in protecting their information privacy, is the "converse of learned helplessness" (Spreitzer, 1995, p. 1444). Formally, Thomas and Velthouse (1990) explained impact through the concept of learned helplessness in humans (Abramson et al., 1978). Learned helplessness is either externally (universal helplessness) or internally (personal helplessness) focused (Abramson et al., 1978). In a privacy context, when consumers experience universal helplessness, they do not see their actions leading to an impact on their privacy protection, regardless of the measures they take or the efforts they exert. This leads to a digital resignation (Draper & Turow, 2019) or "privacy is dead" sentiment (Acquisti et al., 2020), which is characterized as depressed affect and reduced motivation to perform a task (Abramson et al., 1978). Consumers experiencing personal helplessness see making an impact to their privacy protection as possible, but they lack the competence to perform the task successfully. The perception of making an impact is the additive effect of meaningfulness, competence, and awareness of risks that shapes one's attitude on the expectancy of making a difference through one's actions.

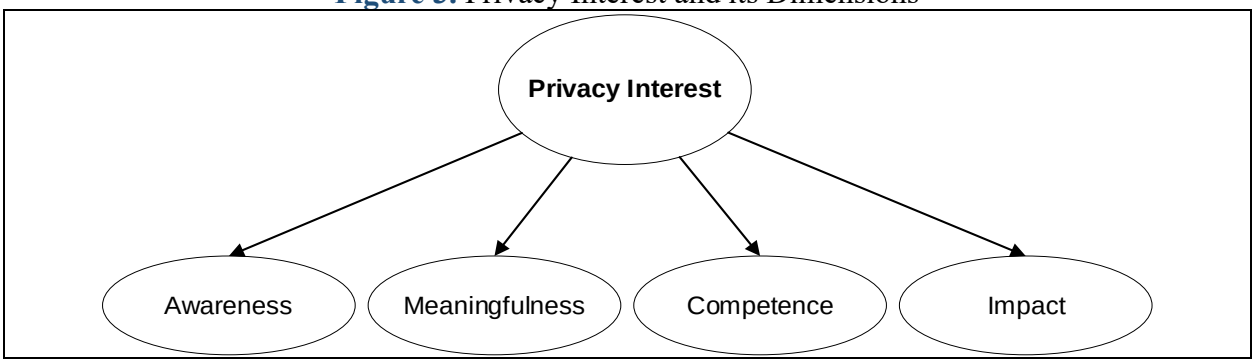
2.3.4. Privacy Interest: A Multidimensional Construct

In summary, we ground our conceptualization of privacy interest on the cognitive model of empowerment, namely the reciprocal relationship between task and global assessments, and layer it with the four stages of interest development to explain consumers' intrinsic task motivation to

improve their information privacy posture. Awareness of a privacy event is the situational trigger that catalyzes a consumer to think of their privacy and its value to them. In time, consumers who develop a privacy interest will improve their privacy competency to enhance the impact they can have on their information privacy protection. However, others who are aware of privacy events and who may also value their privacy may not develop a greater interest in furthering their privacy competency or in learning ways to improve their impact on privacy protection. The four-dimensional conceptualization of privacy interest effectively captures the variance among those with low, medium, and high-levels of interest in privacy protection and can also be used as a proxy of privacy attitude, serving as a functional alternative to privacy concern in contexts leading to the privacy paradox.

Formally, we conceptualize privacy interest as a second-order construct that is measured reflectively. The underlying theoretical theme that undergirds privacy interest is the various stages of interest development from which situational interest becomes dispositional interest. In accordance with our theorization, privacy interest reflects the different stages of interest development, which will manifest in the set of four cognitions—with those who are more dispositionally interested in privacy to perceive higher levels of competency and ability to have an impact on their information privacy, and with those who are low in interest to remain only situationally interested in their privacy and its protection. **Figure 3** illustrates the reflective structure of privacy interest and its correlated subdimensions.

Figure 3. Privacy Interest and its Dimensions



2.4. Scale Development and Validation

The extant privacy research discourse commonly uses privacy concerns as a measurable proxy for the perception of privacy (Smith et al., 2011). We posit that privacy interest can also serve as a proxy for privacy. We propose a privacy interest measurement scale and develop it in accordance with the scale development procedures recommended by MacKenzie et al. (2011) and Weber (2021). **Table 2** describes the steps taken in the scale development process.

Table 2. Overview of Scale Development Procedures

Scale development steps (MacKenzie et al., 2011; Weber, 2021)		**Details of the steps performed in this study**
Conceptualization	Step 1: Develop a conceptual definition of the construct	Adapted privacy interest from student interest (Frymier et al., 1996) and empowerment (Schultz & Shulman, 1993; Spreitzer, 1995; Thomas & Velthouse, 1990) that includes awareness, competence, impact, and meaningfulness as first-order dimensions.
Development of Measures	Step 2: Generate items to represent the construct	Adapted privacy interest items from existing scales and authors' "deduction from the theoretical definition of the construct" (MacKenzie et al., 2011, p. 304).
	Step 3: Assess the content validity of the items	Checked content validity with Amazon MTurk workers using the "pick-and-choose" matrix approach described by Anderson and Gerbing (1991), where workers assigned items to the most appropriate subdimension definition (n = 66).

Model Specification	Step 4: Formally specify the measurement model	Specified privacy interest as a second-order reflective construct. In AMOS, our measurement model includes one path fixed at 1.0 between the latent construct and each of its indicators.
Scale Evaluation and Refinement	Step 5: Collect data to conduct pretest	Performed an initial pretest (two-wave survey) of the psychometric properties with Prolific participants to evaluate the scale's convergent, discriminant, and nomological validity and reliability (n = 423).
	Step 6: Scale purification and refinement	Used AMOS and SPSS to evaluate the goodness of fit of the measurement model, assess the validity and reliability of the set of indicators, revise scales using confirmatory factor analysis, and assess nomological validity.

2.4.1. Generate Items that Represent the Construct

First, we generated items that represented the privacy interest construct. We began by compiling a list of items from an established and validated student interest scale used in the empirical context of classroom learning to measure students' choice, meaningfulness, impact, and competence dimensions. We then performed a qualitative survey with 198 technology users to understand whether any distinctions among the subdimensions in the student interest, privacy concern, and privacy risk concepts would emerge. We asked respondents to rank order the privacy concepts that first came to mind when using their tech devices or visiting a website, and to explain the rationale behind their chosen order. We open-coded the responses to understand the dimensions that emerged in users' privacy considerations. Finally, we reconciled the dimensions of student interest with our theoretical description of privacy interest to identify the relevant dimensions inherent in student interest that would apply to a privacy context. We observed subdimensions related to privacy concerns such as access, secondary use, surveillance, and more. We also observed subdimensions related to interest.

A few illustrative examples of the interest subdimensions we observed in respondents' comments are as follows:

- *(Lack of) Choice*: “I work in data analytics and know how data can be used. I just hope the people collecting the data will treat it ethically and do the right thing.”
- *Awareness and experience*: “I was a computer programmer for 40+ years before I retired. I know how bad security is on all platforms including even the largest computers. Phones are the worst. I do not do any financial apps on my phone, nor do I pay any bills that require my personal info to be entered.”
- *Meaningfulness*: “My personal data is definitely the thing I think about most. Mostly because it's the thing I hear about being compromised the most. I've had it happen to me in the past.”
- *Competence*: “If you don't have interest in protecting your data you won't be able to be proactive when going on websites. You have to know and understand the risks in order to be concerned about the risks.”
- *Impact*: “Risk is important but is not always present within my mind. If I am worried about risk, then the concern would naturally follow. My main focus is interest as having a constant protection will prevent the other two items.”

After evaluating the open codes and the dimensions of student interest, we determined that awareness, meaningfulness, competence, and impact were dimensions directly applicable to a privacy context. Choice, however, was excluded because many of the respondents indicated that they did not have a choice in whether to share their data or to express how their data should be used. We thus omitted choice from the conceptualization of privacy interest.

2.4.2. Assessing the Content Validity of Items

Our second step was to ensure that the items we created for each measure corresponded to the conceptual definition of the respective dimension. We conducted a qualitative item-matching activity to assess the content validity of 40 items included in the item pool. The 40 items were carefully reworded from the existing student interest and learner empowerment scales. We then collected data from Amazon Mechanical Turk (MTurk) respondents (n = 66) who reviewed a matrix of 40 items (rows) and four construct definitions (columns). They were tasked with selecting the most appropriate construct definition to which each item conceptually represented. As a quality check, participants were asked to explain “why” they matched items with specific construct definitions.

We computed two coefficients to measure the content validity of the respondents’ item-

matching activity: (1) proportion of substantive agreement and (2) substantive validity coefficient. Appendix 2E shows the content validity survey instruments and results. A summary of the results is as follows: (1) Choice: five validated items out of nine total items; (2) Competence: nine validated items out of nine total items; (3) Impact: three validated items out of 11 total items; (4) Meaningfulness: five validated items out of 11 total items.

Next, based on the results, we revised several items to align the item wording with the conceptual definition. Specifically, two of the authors conducted iterative rounds of review and assessment of the newly revised items. Each author created a pool of revised items, then discussed the aim of each revision, explaining how and why they revised each item. The authors identified any confusing or ambiguous items and removed them from the pool. After revising the items through multiple iterations of wording revisions and discussions, we retained a total of 20 items in the item pool—five items for each respective interest dimension.

2.4.3. Pilot Study

For our third step, we conducted a pilot study on Amazon Mechanical Turk with 104 respondents. Our objective was to quantitively validate the measurement scales, determine which items to trim, and enhance the clarity of the item wording. Respondents were paid $0.60 USD for their participation. We presented respondents with an infographic, entitled, “Your identity is a steal on the Dark Web” (Experian) that showed the market rates of consumers’ data sold on the Dark Web. We then asked respondents to answer a series of items regarding their willingness to sign up for a new online financial service account, privacy interest subdimensions, privacy concern subdimensions, privacy risk, global information privacy concern, and a realism check (i.e., “There is a high likelihood that consumer data are available for sale on the dark web.”). We performed an exploratory factor analysis and reliability analysis and reviewed the performance of each privacy-related predictor on self-disclosure as a proof-of-concept check.

Cronbach's alpha and composite reliability indicated good reliability for each subdimension of interest within acceptable limits, ranging from 0.91 (choice as lowest) to 0.97 (competence as highest). The average variance extracted (AVE) for meaningfulness (0.63), impact (0.56), and competence (0.81) were greater than 0.50 (Fornell & Larcker, 1981), indicating adequate convergent validity. However, the AVE for choice was below acceptable limits of adequate convergent validity at 0.38. Adequate discriminant validity was also observed for meaningfulness, impact, and competence, with self-loadings significantly higher than any cross-loadings on respective factors. But similar to results obtained in the student interest scale (Frymier et al., 1996), we observed that choice items cross-loaded onto impact, indicating inadequate discriminant validity with choice, which thereby substantiated our decision to replace choice with awareness.

Following previous interest scale development studies (e.g., Frymier et al., 1996; Weber & Patterson, 2000), we removed choice as a subdimension to interest. The possible inapplicability of choice in a privacy decision-making context was confirmed through our qualitative and quantitative assessments in this scale development. We refined item wording and trimmed the number of interest subdimension items from 10 items on average to five items per dimension. The items included in our measurement scales for the main study are shown in Appendix B.

2.4.4. Examining Scale Properties with a New Sample

The fourth step was to purify and refine our newly developed measurement scales (MacKenzie et al., 2011; Weber, 2021). To do so, we performed a two-wave data collection (Podsakoff et al., 2003) with survey respondents using the Prolific crowdsourcing platform. We used Qualtrics to deploy our online surveys.

In the first wave, we deployed a questionnaire to collect survey responses on the following independent and control variables: privacy interest subdimensions (self-developed); internet privacy concern subdimensions, privacy risk, trusting beliefs (Hong & Thong, 2013); perceived

benefits (Kim et al., 2008); global information privacy concern, misrepresentation of identification, Internet experience (Malhotra et al., 2004); prior privacy experience Xu et al. (2012a); and blue attitude marker variable (Miller & Chiodo, 2008). We included response set items to ensure participants were paying attention to the questionnaire; respondents who failed to answer these correctly were automatically screened for removal by the Qualtrics system. Respondents who successfully and validly completed the survey were invited to participate in the second wave data collection.

We administered the second questionnaire, two weeks after the initial survey deployment, to collect survey responses from the wave 1 respondents on the following dependent variables and demographic information: self-disclosure (Jiang et al., 2013); removal from company database, complaining directly to online companies, complaining indirectly to third-party organizations, misrepresentation of personal information (Son & Kim, 2008); blue attitude marker variable (Miller & Chiodo, 2008); global interest in information privacy rating scale, open comment on interest in privacy (self-developed); and demographic information (gender, education, income, employment status, ethnicity). We included attention checks to screen for inattentive respondents. In total, 423 out of 500 Prolific respondents completed waves 1 and 2 questionnaires for a response rate of 84.6% (45% females; age mean = 42.89, with a range between 19 and 82 years; 97.9% with more than 7 years of internet experience; 87.8% misrepresent their identity less than half of the time). The construct means scores for participants who completed waves 1 and 2 surveys were not significantly different from those who completed wave 1 only. Respondents were paid the equivalent of $8.50 per hour, with each survey taking about 7.5 minutes to complete. To assess the psychometric properties of the measurement scales, we used the covariance-based structural equation modeling approach in AMOS v28.

2.4.5. Scale Purification and Validation Process

Next, we describe the scale purification and validation process and results. We followed the guidelines by MacKenzie et al. (2011) and assessed the validity and reliability of the individual indicators and the sets of indicators at the construct level. We performed goodness of fit assessments on the measurement model and examined the convergent validity, discriminant validity, and reliability for privacy interest in a first-order, disaggregated form and a second-order factor structure.

After a scale is established, a confirmatory factor analysis (CFA) is required to assess the convergent validity, discriminant validity, and reliability of the measures (Hair et al., 2012). Accordingly, we conducted a CFA and established that the privacy interest scale exhibited convergent validity (AVEs > 0.5) and reliability (Cronbach's alpha and composite reliability > 0.70) at levels above the acceptable limits (Fornell & Larcker, 1981; Hair et al., 2011). To assess discriminant validity, we checked that all items loaded onto their respective constructs above 0.707 (Straub et al., 2004). For items cross-loading onto other constructs, we followed the recommendation by Gefen and Straub (2005) that cross-loading differences should exceed 0.1. Moreover, we examined the square roots of the AVEs to ensure they exceeded the intercorrelations between constructs. **Table 3** reports the interconstruct correlations. **Table 4** shows the factor loadings and validity and reliability measures and **Table 5** shows the CFA fit indices for the constructs we tested in our measurement and structural models.

Table 3. Interconstruct Correlations

	01	02	03	04	05	06	07	08	09	10
01 PRI	**0.784**									
02 IMP	-0.019	**0.903**								
03 MNG	0.218	0.354	**0.890**							
04 CMP	-0.042	0.619	0.322	**0.900**						
05 AWA	0.296	0.215	0.662	0.218	**0.916**					
06 COL	0.385	0.184	0.658	0.074	0.762	**0.786**				

07 USE	0.376	0.093	0.507	0.028	0.679	0.854	**0.937**			
08 ERR	0.237	0.104	0.397	0.101	0.459	0.484	0.461	**0.912**		
09 ACC	0.443	0.031	0.464	-0.029	0.581	0.769	0.745	0.493	**0.926**	
10 CTL	0.385	0.232	0.664	0.166	0.800	0.861	0.762	0.416	0.638	**0.900**
11 RB	0.372	0.096	0.469	-0.025	0.555	0.706	0.582	0.469	0.573	0.598
12 TR	-0.313	0.342	-0.046	0.321	-0.263	-0.353	-0.392	-0.144	-0.469	-0.272
13 BN	-0.011	0.249	0.187	0.165	0.110	0.110	0.087	0.051	0.094	0.154
14 MRK	-0.011	0.159	0.189	0.166	0.124	0.077	0.115	0.072	0.095	0.059
15 SD	-0.088	-0.008	-0.209	-0.002	-0.289	-0.322	-0.244	-0.105	-0.221	-0.300
16 REM	0.239	0.155	0.432	0.070	0.497	0.483	0.395	0.249	0.407	0.492
17 COM	0.177	0.151	0.337	0.062	0.334	0.281	0.175	0.186	0.196	0.303
18 IND	0.155	0.092	0.212	0.082	0.266	0.201	0.169	0.128	0.166	0.185
19 MIS	0.174	0.021	0.081	0.085	0.165	0.266	0.203	0.100	0.239	0.172

Table 3. Interconstruct Correlations (Continued)

	11	12	13	14	15	16	17	18	19
11 RB	**0.871**								
12 TR	-0.363	**0.849**							
13 BN	-0.051	0.080	**0.824**						
14 MRK	0.063	-0.013	0.237	**0.807**					
15 SD	-0.315	0.297	0.043	0.052	**0.808**				
16 REM	0.375	-0.125	0.032	0.105	-0.262	**0.942**			
17 COM	0.283	0.004	-0.053	0.151	-0.082	0.590	**0.946**		
18 IND	0.156	0.028	-0.087	0.120	-0.064	0.409	0.575	**0.934**	
19 MIS	0.154	-0.261	-0.024	-0.040	-0.116	0.102	-0.013	0.003	**0.951**

Note. The diagonal elements represent the square root of AVE. ACC = Improper Access; AWA = Awareness; BN = Perceived Benefits; CMP = Competence; COL = Collection; COM = Direct Complaint to Company; CTL = Control; ERR = Errors; IMP = Impact; IND = Indirect Complaint to Third Party Organization; MIS = Misrepresentation ; MNG = Meaningfulness; MRK = Blue Marker; PRI = Prior Experience; RB = Risk Beliefs; REM = Removal from Company Database; SD = Self-disclosure; TR = Trust Beliefs; USE = Secondary Usage

Table 4. Confirmatory Factor Analysis Results (n = 423)

Items	**Mean**	**Standard Deviation**	**Factor Loading**	**Squared multiple correlation**
Privacy Interest				
Awareness (C.A. = 0.939; C.R. = 0.940; AVE = 0.838)				
AWA1	5.48	1.35	0.896	0.799
AWA2	5.59	1.33	0.916	0.844
AWA3	5.67	1.30	0.934	0.872
Meaningfulness (C.A. = 0.952; C.R. = 0.950; AVE = 0.792)				
MNG1	5.46	1.21	0.858	0.715
MNG2	5.71	1.13	0.922	0.845
MNG3	5.91	1.01	0.893	0.834
MNG4	5.71	1.12	0.908	0.842
MNG5	5.92	1.03	0.869	0.786
Impact (C.A. = 0.958; C.R. = 0.956; AVE = 0.815)				
IMP1	5.00	1.29	0.895	0.790

IMP2	4.87	1.28	0.868	0.787
IMP3	4.91	1.30	0.916	0.831
IMP4	4.87	1.29	0.898	0.834
IMP5	4.85	1.29	0.934	0.864
Competence (C.A. = 0.955; C.R. = 0.955; AVE = 0.810)				
CMP1	4.88	1.24	0.892	0.794
CMP2	4.50	1.39	0.875	0.763
CMP3	4.77	1.32	0.929	0.865
CMP4	4.74	1.29	0.912	0.830
CMP5	4.76	1.33	0.891	0.798
Privacy Concern (Awareness Dimension Above)				
Collection (C.A. = 0.827; C.R. = 0.829; AVE = 0.618)				
COL1	5.39	1.32	0.759	0.570
COL2	5.86	1.12	0.745	0.556
COL3	5.68	1.23	0.850	0.726
Secondary Usage (C.A. = 0.955; C.R. = 0.956; AVE = 0.878)				
USE1	5.74	1.19	0.919	0.842
USE2	5.84	1.23	0.944	0.892
USE3	5.86	1.17	0.948	0.901
Errors (C.A. = 0.936; C.R. = 0.937; AVE = 0.832)				
ERR1	4.8	1.56	0.939	0.880
ERR2	4.83	1.54	0.860	0.740
ERR3	4.72	1.56	0.936	0.877
Improper Access (C.A. = 0.947; C.R. = 0.947; AVE = 0.857)				
ACC1	5.69	1.19	0.928	0.860
ACC2	5.59	1.23	0.930	0.868
ACC3	5.65	1.21	0.919	0.841
Control (C.A. = 0.927; C.R. = 0.928; AVE = 0.810)				
CTL1	5.63	1.24	0.914	0.833
CTL2	5.63	1.18	0.920	0.846
CTL3	5.59	1.21	0.866	0.752
Privacy Calculus Constructs				
Risk Beliefs (C.A. = 0.926; C.R. = 0.926; AVE = 0.758)				
RB1	5.01	1.25	0.888	0.795
RB2	4.86	1.32	0.864	0.734
RB3	4.83	1.35	0.886	0.790
RB4	4.61	1.32	0.844	0.711
Trust Beliefs (C.A. = 0.906; C.R. = 0.911; AVE = 0.720)				
TR1	3.65	1.46	0.941	0.896
TR2	3.32	1.57	0.855	0.724
TR3	3.88	1.37	0.876	0.765
TR4	3.86	1.52	0.705	0.492
Perceived Benefits (C.A. = 0.900; C.R. = 0.913; AVE = 0.679)				
BN1	6.39	0.73	0.796	0.633
BN2	5.78	1.12	0.683	0.467
BN3	6.22	0.82	0.894	0.800
BN4	6.22	0.81	0.872	0.759
BN5	6.17	0.88	0.857	0.733
Dependent Variables (Measured in Separate Nomological Networks)				
Self-disclosure (C.A. = 0.903; C.R. = 0.904; AVE = 0.653)				
SD1	2.98	1.51	0.871	0.753
SD2	2.46	1.40	0.785	0.612
SD3	3.61	1.60	0.848	0.718

SD4	3.21	1.52	0.820	0.672
SD5	2.52	1.45	0.707	0.493
Removal from Company Database (C.A. = 0.958; C.R. = 0.959; AVE = 0.888)				
REM1	5.23	1.69	0.976	0.948
REM2	5.24	1.72	0.966	0.931
REM3	5.52	1.55	0.882	0.769
Direct Complaint to Company (C.A. = 0.961; C.R. = 0.963; AVE = 0.896)				
COM1	3.96	2.01	0.974	0.941
COM2	4.04	2.02	0.985	0.974
COM3	4.47	1.93	0.877	0.762
Indirect Complaint to Third Party Organization (C.A. = 0.951; C.R. = 0.953; AVE = 0.872)				
IND1	3.10	1.89	0.966	0.933
IND2	3.20	1.90	0.975	0.949
IND3	3.75	1.99	0.856	0.732
Misrepresentation (C.A. = 0.965; C.R. = 0.966; AVE = 0.905)				
MIS1	3.78	2.01	0.972	0.946
MIS2	3.82	1.97	0.991	0.980
MIS3	4.24	1.97	0.888	0.790
Antecedent				
Prior Experience (C.A. = 0.802; C.R. = 0.821; AVE = 0.614)				
PRI1	3.14	1.36	0.856	0.730
PRI2	4.28	1.31	0.550	0.309
PRI3	2.86	1.25	0.899	0.800
Marker Variable				
Blue Marker (C.A. = 0.906; C.R. = 0.926; AVE = 0.652)				
MRK1	4.09	1.78	0.578	0.334
MRK2	5.17	1.43	0.653	0.426
MRK3	4.59	1.48	0.560	0.314
MRK4	6.03	0.92	0.954	0.911
MRK5	5.96	0.97	0.920	0.847
MRK6	5.96	0.97	0.961	0.923
MRK7	6.06	0.86	0.899	0.809

Note. AVE = Average variance extracted; C.A. = Cronbach's alpha; C.R. = Composite reliability

Table 5. CFA Fit Indices for Model (n = 423)

Model fit	**Recommended value**	**Measurement model**	**Measurement model with common method factor**
$\chi 2$	N/A	4454.427	3852.030
Df	N/A	2380	2307
$\chi 2$ /df	≤ 5	1.872	1.670
SRMR	< 0.10	0.039	0.037
GFI	≥ 0.90	0.771	0.804
AGFI	≥ 0.80	0.740	0.770
NFI	≥ 0.90	0.876	0.893
CFI	≥ 0.90	0.938	0.953
RMSEA	≤ 0.08	0.045	0.040
AIC	N/A	5096.427	4640.030
CAIC	N/A	6716.633	6628.695

We found that the collection dimension in privacy concern had discriminant validity issues. Namely, the square root of the AVE for collection was less than the absolute value of its correlation

with secondary usage and with control. The collection dimension was highly correlated with other dimensions in privacy concern as observed in its cross-loadings. It could be that users who are concerned about restricting control were similarly concerned about the collection and unauthorized secondary use of their data beyond what they can feasibly control. We used the privacy concern scale verbatim in its original form without any adjustments to fit a specific context. This is because we were assessing dispositional privacy interest and privacy concern attitudes without the use of a contextual stimulus (situational trigger) to instigate situational privacy attitudes. Because privacy concern was measured reflectively, we expected that the dimensions would covary. The original privacy measurement scales by (Hong & Thong, 2013) and (Malhotra et al., 2004) showed high corrections between the first-order concern dimensions. The newly developed privacy interest scale did not have any validity issues, and the privacy concern scale was used as a baseline to compare with privacy interest.

Further, we assessed whether common method bias (CMB) was an issue in our two-wave data collection. In our CFA model, we included a latent common method factor that loaded on the items for 19 first-order constructs listed in **Table 3** (independent, dependent, control, and marker variables). We evaluated the zero-constrained and unconstrainted models and performed a chi-squared significance test of each model's fit indexes to assess whether the models were the same, or invariant. We were unable to reject the null hypothesis (i.e., both models are the same), which indicated that CMB was likely not an issue in our assessments. However, we retained a marker variable in our nomological validity assessments as a control in our structural models.

2.4.6. Measurement Models for Privacy Interest

To examine the privacy interest scale, we first empirically evaluated privacy interest in its first-order disaggregated form. We then evaluated privacy interest in its proposed second-order factor structure. The next step was to assess the validity of the disaggregated and second-order factor

structures and formally compare them. We followed existing privacy research on scale development and compared the model fit of the theoretical second-order factor of privacy interest to the model fit of its disaggregated first-order form. **Table 6** shows the fit indices for the disaggregated and the higher-order factor structures, and **Figure 4** shows each of the models in their respective forms. To compare the good-of-fit indices between the two models, Marsh and Hocevar (1985) recommended calculating a target coefficient (*t*-value). The target coefficient approach has been used to validate previous privacy concern scales that examined disaggregated versus higher-order forms (e.g., Hong & Thong, 2013; Stewart & Segars, 2002). Formally, a target coefficient is scaled from 0 to 1 and is calculated by dividing the model fit ($\chi 2$) of the disaggregated model by the model fit ($\chi 2$) of the more restrictive higher-order model. A target coefficient approaching the upper limit of 1 suggests that the correlation among the first-order indicators can be fully explained by the higher-order factor structure (Marsh & Hocevar, 1985).

We calculated the target coefficient by setting the model fit ($\chi 2$) of the disaggregated model as the dividend (393.489) and the model fit ($\chi 2$) of the higher-order model (557.935) as the divisor to obtain a *t*-value of 0.705. A larger *t*-value indicates a better model. This provides reasonable support that the theorized second-order factor structure of privacy interest provides an adequate explanation of the correlations among the first-order indicators.

Figure 4. Assessing the Higher-Order Factor Structure for Privacy Interest

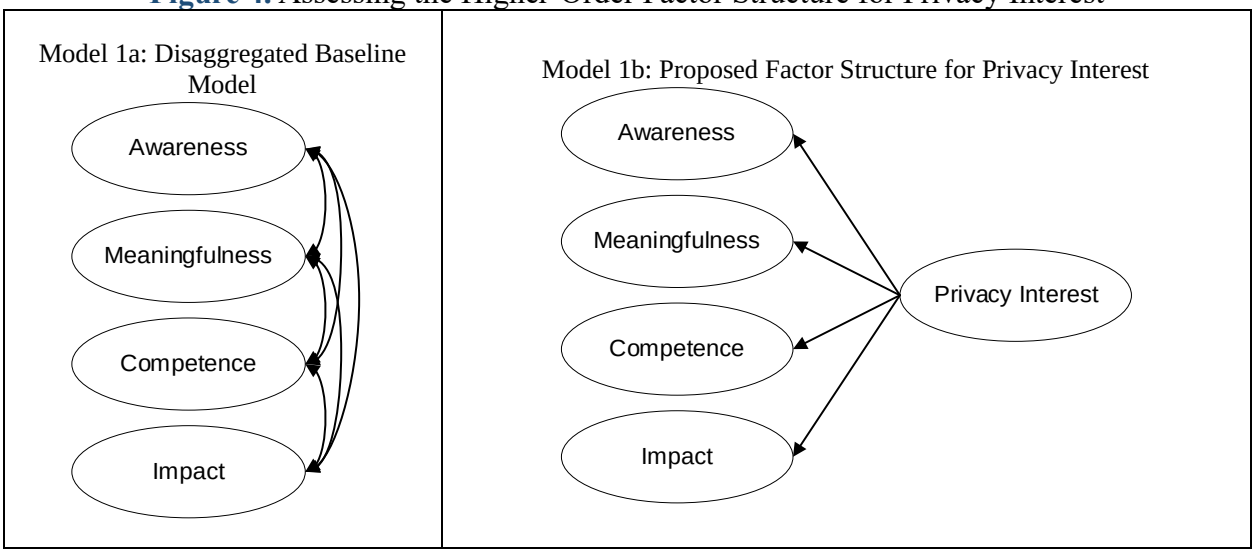


Table 6. CFA Fit Indices for Measurement Models

Model fit	**Recommended value**	**Model 1a (disaggregated)**	**Model 1b (privacy interest)**
χ 2	N/A	393.489	557.935
Df	N/A	129	131
χ 2 /df	≤ 5	3.050	4.259
SRMR	< 0.10	0.030	0.127
NFI	≥ 0.90	0.955	0.936
CFI	≥ 0.90	0.969	0.950
RMSEA	≤ 0.08	0.070	0.088
AIC	N/A	513.489	673.935

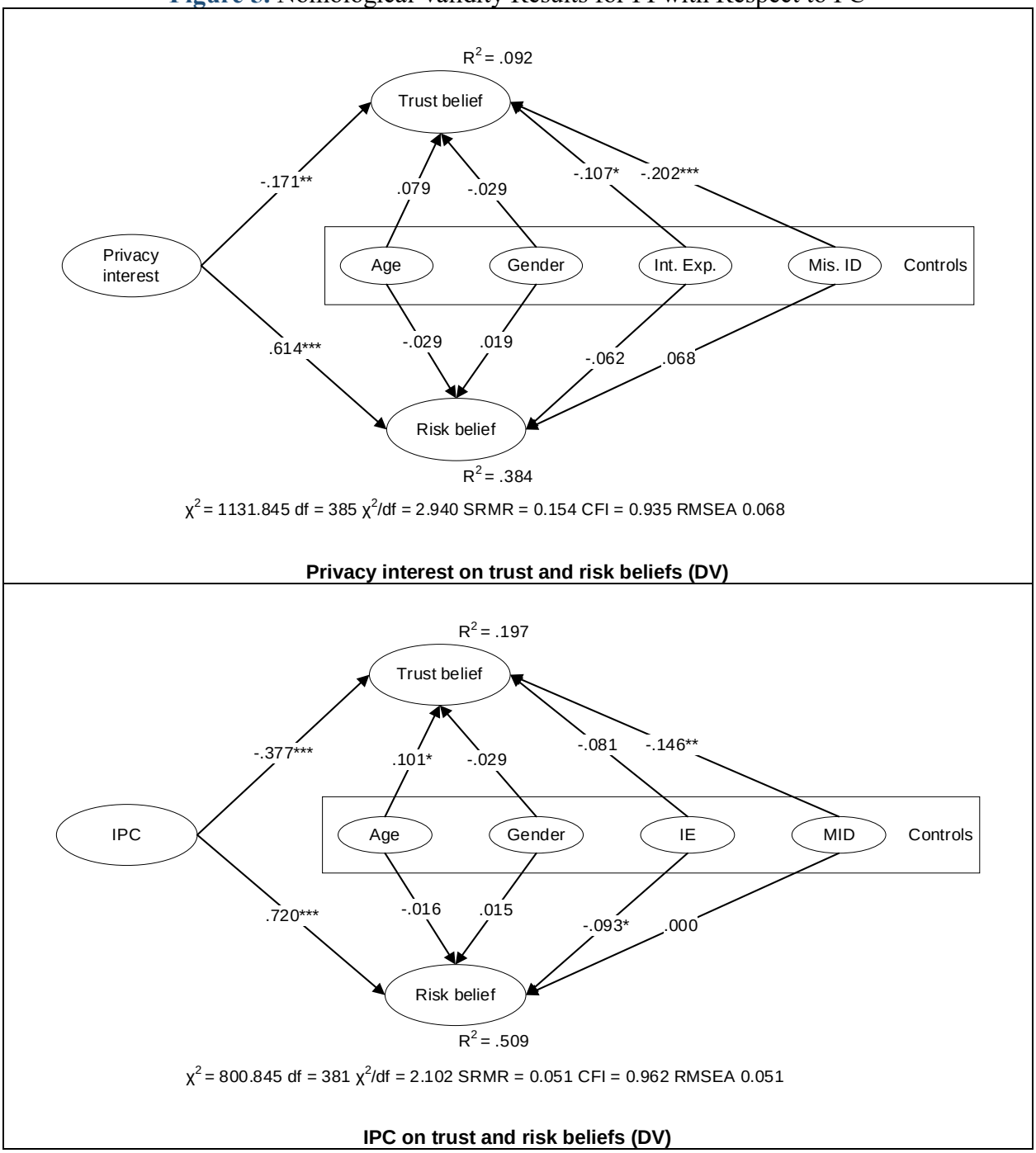
2.4.7. Assessing Scale Validity

After confirming that the privacy interest scale meets acceptable ranges of good psychometric properties, we assessed the scale validity within multiple nomological networks of constructs, referred to as nomological validity (MacKenzie et al., 2011). *Nomological validity* refers to "whether the indicators of the focal construct relate to the measures of other constructs" in a theoretical network (MacKenzie et al., 2011, p. 317). We modeled nomological networks used in Hong and Thong (2013)— specifically testing whether privacy interest is significantly related to other constructs typically hypothesized within a privacy concern-based nomological network (i.e., risk beliefs, trusting beliefs). We also followed an approach similar to Son and Kim (2008) and tested the focal construct in separate nomological networks of constructs with different outcome

variables (i.e., self-disclosure, removal, direct complaint, indirect complaint, misrepresentation). Finally, we self-developed a global interest in information privacy rating scale (1 to 100) and tested whether a significant relationship existed with the theorized privacy interest.

First, we evaluated the effect of privacy interest on trusting belief and risk belief. Hong and Thong (2013) and Malhotra et al. (2004) examined their self-developed privacy scales by examining the relationship of IPC and IUIPC, respectively, on trust and risk beliefs. Accordingly, by positing that privacy interest can serve a proxy role similar to privacy concern, we expected privacy interest to have a significant negative relationship with trusting belief and a significant positive relationship with risk belief. We also included control variables (age, gender, internet experience, misrepresentation of identity) that research has shown to possibly influence trusting and risk beliefs (Jiang et al., 2013; Zhang et al., 2022). **Figure 5** shows the results and fit indices for privacy interest- and privacy concern-based structural models. The fit indices for privacy interest are mostly in acceptable limits, which indicate acceptable fit with the data. As expected, the relationship from privacy interest to trusting belief (-0.171) is significant (negative) and the relationship from privacy interest to risk belief (0.614) is significant (positive). This indicates that privacy interest can serve as a satisfactory proxy in a nomological network with risk and trust beliefs, exhibiting good nomological validity. However, the path coefficients for the relationships between privacy concern and trust belief (-0.377) and privacy concern and risk belief (0.720) are stronger than those in the relationships involving privacy interest.

Figure 5. Nomological Validity Results for PI with Respect to PC



Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = misrepresentation of identity

2.4.8. Assessing Nomological Validity

Second, to further validate the psychometric properties of privacy interest, we developed two scales using the six dimensions of privacy concerns validated by Hong and Thong (2013): (1)

Internet privacy concern comprising collection, secondary usage, errors, improper access, control, and awareness (Hong & Thong, 2013), and (2) CFIP comprising collection, errors, secondary usage, and improper access (Smith et al., 1996). We included each of the privacy concern constructs in separate structural models, along with privacy interest, privacy risk, trust, benefits, and one of five protective-privacy behavioral intentions to assess nomological validity. Previous studies have tested and validated the effect of privacy concern on the following five protective-privacy responses:

- *Direct complaint* is a public action where customers complain directly to the company in question (Son & Kim, 2008).
- *Indirect complaint* is a public action where customers who are dissatisfied by the redress provided by the company in question proceed to take action through a third-party organization (Son & Kim, 2008).
- *Removal* is a private action where customers request the removal of their personal information from a company's databases (Son & Kim, 2008).
- *Self-disclosure* is the provision of true personal information of oneself (Jiang et al., 2013).
- *Misrepresentation* is the provision of false personal information of oneself (Jiang et al., 2013).

Our primary focus was to evaluate the predictive ability of privacy interest on each of the five behavioral intentions with and without the two privacy concern constructs in the nomology. We used the APCO macromodel as a well-tested theory in extant information privacy research to empirically estimate five (5) structural models to test the comparative effects of privacy interest and privacy concern on a protective-privacy response: (1) direct complaint, (2) indirect complaint, (3) removal, (4) self-disclosure, and (5) misrepresentation. Model "A" included privacy interest in the APCO without privacy concern (IPC). Model "B" included privacy concern (IPC) in the APCO without privacy interest. Model "C" included privacy concern (IPC) as an imputed factor score variable and privacy interest as a latent variable in the APCO. Model "D" included privacy concern (IPC) as a latent variable and privacy interest as an imputed factor score variable in the APCO.

Model “E” included privacy concern (CFIP) and privacy interest in the APCO.

For each model, we included four control variables (age, gender, internet experience, misrepresentation of identity) and the blue marker variable. Also, we included prior experience as an antecedent to privacy concern (Dinev et al., 2015; Smith et al., 2011). Notably, structural models including IPC and privacy interest as latent variables could not be estimated because each construct shared “awareness” as a subdimension. For comparison purposes only, we imputed a factor score for each subdimension of IPC and privacy interest and estimated models with one of each included (i.e., Model C and Model D). As a result, the goodness-of-fit measures for Models C and D did not meet acceptable limits of good model fit. Therefore, we did not draw any inferences from the results obtained in Models C and D and provided them only for illustrative purposes. Instead, when we made direct comparisons between privacy concern and privacy interest in the same model, we estimated CFIP, which did not include the “awareness” dimension, and therefore drew inferences from Model “E.”

(1) Direct complaint: For the first comparison, we compared the privacy interest scale to the IPC (Hong & Thong, 2013) and CFIP (Smith et al., 1996) scales in predicting direct complaint intentions. The purpose behind this comparison was to demonstrate privacy interest as a viable alternative to privacy concern in predicting consumers’ public action (Son & Kim, 2008) and mobilization (Leidner & Tona, 2021) protective privacy behavioral intentions. Examples of public action privacy-protective responses include complaining directly to online companies and indirectly to third-party organizations (Son & Kim, 2008). Leidner and Tona (2021) described mobilization responses as those where consumers share a collective concern toward a companies’ misuse of personal data and then attempt to exert pressure on the company or against the data practice. Path coefficients, R^2 values, and model fit indices for the five structural models are shown

in Appendix 2D, **Figure D1**. The results showed that the privacy interest model (1A) explained more variance in direct complaint than the IPC model (1B; 19.6% vs. 16.6%) and that privacy interest was a significant predictor of direct complaint, whereas CFIP was not (1E; path coefficient: 0.348 vs. -0.057), suggesting privacy interest to be the more effective predictor of direct complaint than the two types of privacy concern.

(2) Indirect complaint: The second public action privacy-protective response was complaining to third-party organizations (Son & Kim, 2008). Results (Appendix 2D, **Figure D2**) showed that the privacy interest and IPC models explained approximately the same level of variance in indirect complaint (13.4% vs. 13.6%). The path coefficients between privacy interest and indirect complaint and between IPC and indirect complaint were each significant (p-value < 0.001; 0.266 vs. 0.259). However, when compared with CFIP, the effect of privacy interest was higher than the effect of CFIP (path coefficients: 0.221 vs. 0.104) on indirect complaint. Generally, the results suggested that privacy interest was an equally effective predictor as IPC and was more effective than CFIP in predicting consumers' public action, mobilization privacy responses. The next three protective-privacy responses are considered micro-level responses, which are actions taken primarily by a single consumer (Leidner & Tona, 2021).

(3) Removal: Removal is a private action consumers take to remove their personal information from a company's database, such as from personalized marketing campaigns, which can impact a company's ability to develop customer loyalty programs (Smith et al., 1996; Son & Kim, 2008). Similar to the results for indirect complaint, the privacy interest and IPC models (Appendix 2D, **Figure D3**) explained approximately the same level of variance in removal (27.7% vs. 27.1%); the path coefficients between privacy interest and removal and between IPC and removal were each significant (p-value < 0.001; 0.466 vs. 0.494); and the effect of privacy interest was higher than

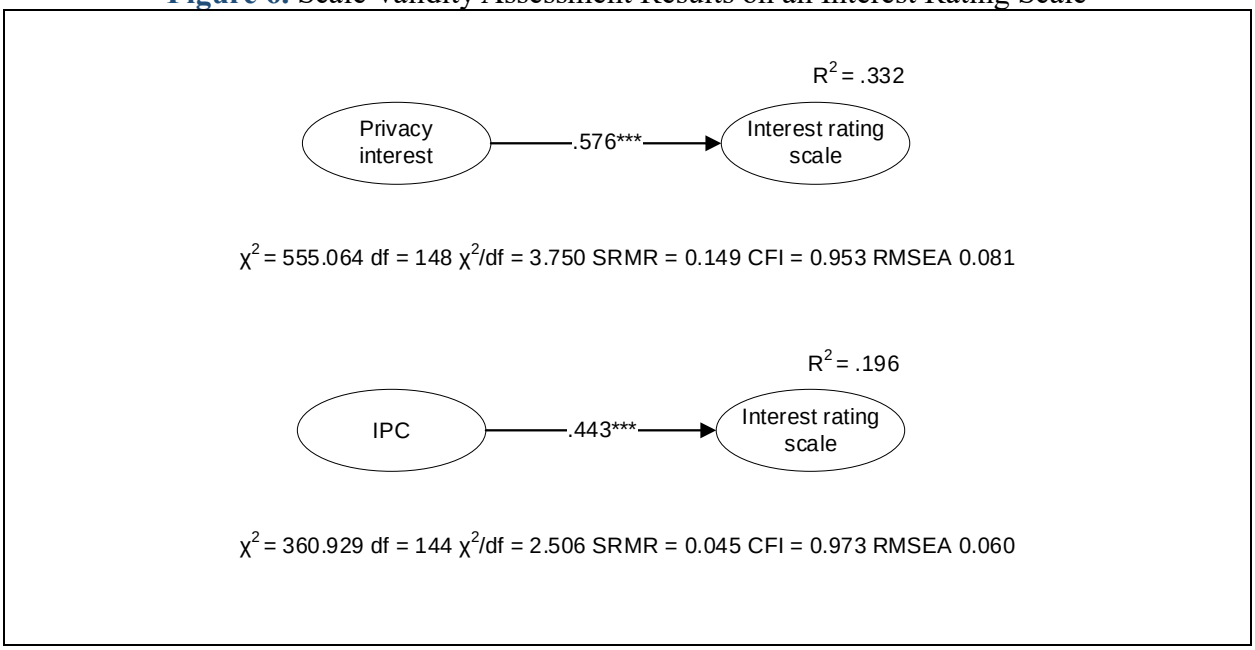
the effect of CFIP (path coefficients: 0.375 vs. 0.202) on removal.

(4) Self-disclosure: Self-disclosure occurs when a consumer gives *true* personal information to a company (Jiang et al., 2013). Results (Appendix 2D, **Figure D4**) showed that the privacy interest model (A) explained more variance in self-disclosure than the IPC model (B; 18.6% vs. 16.4%). Moreover, in comparison to CFIP, privacy interest was a significant predictor (path coefficient: -0.191 vs. 0.079), whereas CFIP was not. In general, the results suggested that privacy interest was a more effective predictor of self-disclosure than CFIP and IPC.

(5) Misrepresentation: In contrast to self-disclosure, misrepresentation occurs when a consumer provides *false* personal information to a company (Jiang et al., 2013). The privacy interest model explained approximately the same level of variance in misrepresentation as the IPC model (32% vs. 31.3%; Appendix 2D, **Figure D5**). However, when CFIP and privacy interest were included in the same model, privacy interest was not a statistically significant predictor of misrepresentation, whereas CFIP was a statistically significant predictor (path coefficient: 0.058 vs. 0.161). We note that the control variable, misrepresentation of identity, was the strongest predictor of misrepresentation behaviors.

(6) Interest rating scale: At the conclusion of our Wave 2 survey, we asked participants to indicate their level of interest in information privacy by responding to the following question: "I see why I should be interested in my information privacy in today's digital age" [1 = Very untrue for me; 100 = Very true for me] (**Figure 6**). We found that privacy interest explained more variance in the interest rating scale than IPC (33.2% vs. 19.6%), although both privacy interest and IPC were statistically significant predictors (path coefficient: p-value < 0.001; 0.576 vs. 0.443).

Figure 6. Scale Validity Assessment Results on an Interest Rating Scale



Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns. Interest rating scale: Indicate your level of interest in information privacy, "I see why I should be interested in my information privacy in today's digital age." [Slider scale: 1 = Very untrue for me; 100 = Very true for me]

2.5. Discussion

Privacy has been an often-discussed topic for greater than a century (e.g., Warren & Brandeis, 1890). The importance of information privacy is heavily engrained in IS research—recognized as one of the most important ethical issues of the information age (Mason, 1986)—and remains a forefront issue to this day (Acquisti et al., 2022). However, as new technologies emerge and societies shift from the information (disclosure and use) age to the autonomous (machine learning) age, research has remained focused on investigating consumers' privacy concerns, which magnified in importance since the call to action by Mason (1986). We assert that information privacy research can benefit from a new theory-driven conceptualization and a validated measurement scale of privacy interest. The premise of current information privacy research is that consumers can appropriately perform a privacy calculus on the benefits and risks of technology use and rationally choose whether to disclose information or use a technology. Smart cities, smart cars, and other smart devices that are owned and deployed by other consumers or organizations

can also collect data without consumers' explicit consent (e.g., smart doorbells capturing citizens' biometric data). By including interest in privacy as part of consumers' privacy calculus, researchers and practitioners can more adeptly understand the limits of rampant data collection and pervasive technologies and the extent to which such factors prompt consumers' mobilization and privacy protective behaviors. Thus, focusing on interest rather than concern can illuminate a new trajectory of privacy research that is focused on consumers' proactive rather than reactive behaviors.

Accordingly, we draw on the cognitive model of empowerment (Thomas & Velthouse, 1990) and stages of interest development (Hidi & Renninger, 2006) to propose a multidimensional conceptualization of privacy interest. Privacy interest encompasses situational and dispositional interest dimensions that develop through key phases (Hidi & Renninger, 2006). Phase 1 is triggered situational interest and represents the awareness dimension where a consumer becomes aware of a privacy event, such as Tesla employees sharing customers' in-car camera images (Stecklow et al., 2023), which may elicit a positive or negative feeling in the person. Phase 2 is maintained situational interest, where a consumer develops a sense of value toward preserving their privacy in the triggered privacy event, which compels the person to assess the meaningfulness of their information privacy. Phase 3 is emerging individual (dispositional) interest and occurs when a consumer has stored knowledge (competence) and value (meaningfulness) toward their information privacy and has positive feelings toward privacy protection. Phase 4 is well-developed individual (dispositional) interest, the zenith of privacy interest, and is demonstrated when a consumer perseveres through the challenges of reengaging in protective-privacy behaviors (e.g., continually adjusting privacy settings, removing data from data broker lists, opting out of personalization services), knowing their actions can positively influence their information privacy.

Our conceptualization of privacy interest incorporates the stages of interest development as manifested in the first-order dimensions of cognitive empowerment.

The dimensions of privacy interest and the various forms of privacy concern cover different scopes of privacy and can thus be included as a collective toolkit in future privacy empirical research. Whereas the awareness dimension is common in privacy interest, IPC, and IUIPC, it serves a different role in each conceptualization. In privacy interest, awareness is intended to catalyze an introspective assessment of how a consumer values their information privacy and how they assess their efficacy to effectively protect it . In privacy concern, awareness is externally focused and indicates whether a consumer is aware of companies' data handling and use practices (Malhotra et al., 2004). We are confident that our privacy interest scale can therefore complement existing privacy concern scales and hence capture a greater scope of the concept of information privacy, referred to as scope validity (Weber, 2021). Furthermore, privacy interest can provide additional latitude to researchers by serving as an independent proxy of information privacy that can be incorporated into the risk-based equation of the privacy calculus.

2.5.1. Theoretical Contribution

Our study contributes to the information privacy research discourse in several important ways. First, we reconceptualize the interest construct in the privacy calculus model from a confidence and enticement belief to a risk belief. This reconceptualization pivots the focus away from identifying the key traits or aspects that encourage technology adoption toward a focus on consumers' sense of empowerment to anticipate downstream privacy risks and to mitigate the impact of these risks. This distinction is crucial because information asymmetries between companies and consumers are widening, and novel technologies are being deployed faster than consumer protection policies are enacted. For example, Google's CEO Sundar Pichai talked about the mismatch between the pace at which technology evolves and the pace at which humans can

adapt as societal institutions (Pelley, 2023). Therefore, the burden of information privacy protection is placed primarily on the consumer, and our study examines how equipped consumers are at shouldering the responsibility of information privacy protection. We believe that knowing whether consumers have an interest in privacy is crucial for the expansion of future privacy research and for the development of privacy legislation that is mutually beneficial for those who seek greater privacy protection and for those who seek the affordances provided by novel AI technologies.

Second, the privacy paradox is a phenomenon that has perplexed researchers, policymakers, and managers. Namely, consumers claim they have privacy concerns, but their behaviors betray these concerns. Our study provides a different theoretical lens through which to view information privacy. Instead of focusing on privacy concerns, which are mostly elevated among typical Western consumers (Pew Research Center, 2019), we unveil a novel perspective that privacy concern is not the only appropriate proxy for privacy attitude— especially given that different situational factors can exert unequal influences on attitudes and behaviors. This lends to the idea that interest in privacy is a viable, alternative viewpoint for future privacy research (Mulligan et al., 2020).

As our study conceptually and empirically demonstrates, privacy interest offers stronger explanations of consumers' mobilization privacy-protective behaviors than does privacy concern. In some instances, privacy interest is more effective than privacy concern in explaining micro-level, individual behaviors, such as self-disclosure; however, privacy concern is more effective at explaining misrepresentation behaviors. Of interest to researchers may be the dimensions of privacy interest (meaningfulness, competence, impact) that can be separately theorized and examined in relation to other privacy-related construct or privacy concern dimensions. For

example, meaningfulness shares semantic meaning with disposition to value privacy (Chen et al., 2021b; Karwatzki et al., 2022), which is an antecedent to privacy concern (Dinev et al., 2015; Smith et al., 2011), and can separately contribute to privacy concern research. Another example would be competence which shares semantic meaning with self-efficacy, which is a construct broadly studied in information security management research (e.g., Boss et al., 2015; Chen et al., 2021c; Moody et al., 2018). Accordingly, privacy interest can contribute to motivation-based research in security and privacy research grounded in the extended parallel processing model (Witte, 1992), protection motivation theory (Maddux & Rogers, 1983), technology threat avoidance theory (Liang & Xue, 2009), and health belief model (Rosenstock, 1974).

Third, our study conceptualizes the privacy interest construct, develops a valid measurement scale for privacy interest, and illustrates its effectiveness as a scale and construct in multiple nomological networks of constructs with outcome variables that address a broad range of consumer protective-privacy behaviors. The privacy interest construct and measurement scale thus provide researchers with a foundation on which to build and accumulate knowledge pertaining to consumers' disposition toward privacy, which complements the knowledge on consumers' situational reactions to various privacy events.

Our hope in the development of privacy interest is that researchers will expand our theoretical knowledge on individual-level interest and develop theoretical frameworks and theories that explain consumers' introspective decision-making toward privacy. An example includes advancing understanding of situational triggers that make privacy salient and can improve a consumer's privacy posture beyond the single event that triggered their privacy thoughts. Another example could be to develop strategies that can enhance consumers' dispositional interest in privacy to encourage them to develop competence in privacy protection and confidence in knowing

their actions can influence how their data are collected and used. Finally, researchers can further examine the boundary conditions of privacy interest to explain when examining interest is appropriate and when retaining a focus on concern should remain prominent. We aspire to galvanize privacy researchers' engagement with our conceptualization of privacy interest will galvanize privacy researchers' interest, encouraging them to explore consumer-focused information privacy research in an era where machine learning systems, such as computer vision, large language models, and speech recognition, demand vast amounts of data. This work is pivotal, because too often, consumers unknowingly contribute their voice, image, and text data for these systems, which frequently leads to unforeseen risks and negative consequences (Metz, 2023).

2.5.2. Policy, Managerial, and Practical Implications

Assuming our results will continue to broadly hold, our work infers multiple practical, managerial, and policy implications. First, our conceptualization can provide a new public polling tool for citizens to express their interest in information privacy protection, adding a new perspective to privacy in the 21st century (Kennedy et al., 2023). For example, think tanks such as Pew Research Center that broadly survey the public about privacy issues can investigate consumers' privacy interest as they relate to technologies inside the home, car, and workplace. Notably, polling on privacy concern focuses on consumers' perceptions of companies' data practices, whereas polling on privacy interest can unveil the types of mobilization behaviors consumers may engage in when they perceive their information privacy is under threat beyond their tolerable limits.

Second, because privacy interest comprises situational and dispositional interest, companies can use our conceptualization to distinguish dispositional differences in consumers who are highly protective of their privacy and those who are not protective of their privacy. The multidimensional construct allows companies to explore the meaningfulness, competence, and impact perceptions consumers have, which can help companies determine new features or policies to add to ameliorate

users' privacy interest. For example, developers can assess which features prompt consumers' interest in protecting their privacy as opposed to elicit their privacy concerns to develop a solution amenable to consumers and companies. By nature, humans have a desire for privacy (Acquisti et al., 2022), but humans may not have an interest in protecting their privacy; therefore, understanding this distinction can be helpful to developers when designing products or services.

Third, because consumers are considered rational actors who willfully provide their data and consent to privacy policies and terms of conditions, the privacy paradox is used as justification for regulators not to enact policies that intervene in the consumer-company interaction (Solove, 2021; Westin, 2000). "Even worse, courts and policymakers often fail to recognize privacy interests at all" (Solove, 2015, p. 74). Thus, our conceptualization of privacy interest provides an alternative lens to policymakers through which to understand consumer attitude beyond their compliance to notice and consent frameworks. Consumers may readily consent to terms and conditions, but they can still have an interest in their privacy protection, which may manifest in protective behaviors other than self-disclosure, such as public and privacy actions (Leidner & Tona, 2021; Son & Kim, 2008). As an alternative proxy of privacy, the privacy interest scale can provide a quantitative tool to companies, organizations, and policymakers to aid in fine-tuning privacy settings, educational programs, and policies and to identify those with high privacy interests to provide personalized privacy options at a granular, more selective scale. By contrast, consumers with low privacy interest may not find privacy-first designs or services greatly appealing and may instead want benefits associated with personalization and information disclosure.

2.5.3. Limitations and Future Research

Although we carefully followed the leading guidelines for producing scale development with high content, construct validity, and test-retest validity (MacKenzie et al., 2011; Weber, 2021), by collecting multiple rounds of data and rigorously assessing the reliability and validity of our scale,

our research nonetheless has limitations that open compelling opportunities for future research. First, our conceptualization of privacy interest relied on awareness, and we specifically operationalized it using the awareness scale by Hong and Thong (2013). Our intention was to avoid recontextualizing any items from the original scales we operationalized and to retain the items verbatim. This created multicollinearity issues when we included IPC and privacy interest into the same structural model because IPC and privacy interest included the same awareness measure. We were thus unable to examine IPC and privacy interest in the same structural model. We included each as an imputed factor score variable and as a latent model (see Appendix 2D), but the goodness-of-fit measures did not meet acceptable limits. Consequently, we encourage future research to further investigate the dual effects of IPC (or IUIPC) and privacy interest on other types of protective-privacy behaviors beyond the scope of what this study examined.

As a result of including the awareness dimension into the privacy interest scale without adapting the item wording, we obtained lower goodness-of-fit measures for the higher-order privacy interest scale when compared to a disaggregated model. For example, the target coefficient (*t*-value) obtained for the higher-order factor structure was 0.705. Ideally, a *t*-value greater than 0.90 provides strong support for the use of a higher-order factor structure over a disaggregated model as the appropriate model of a construct.

Second, our study employed a temporally separated, two-wave survey and asked respondents to answer independent and dependent variables items at separate points in time (Podsakoff et al., 2003). As such, we can draw only associational inferences from the data. We modeled our nomological networks according to the extended privacy calculus (Dinev & Hart, 2006) and assessed the nomological validity; however, we only developed the privacy interest construct and did not theorize a formal nomology centered around it. Accordingly, a compelling opportunity is

for privacy researchers to identify plausible antecedents, mediators, and consequences for privacy interest and to propose formal privacy theories using it. This will require further methodological exploration and testing, such as through experiments to identify moderators and through qualitative methods to discover constructs and dimensions that may complement privacy interest.

2.6. Conclusion

Relying on the cognitive model of empowerment and the four stages of interest development, we conceptualized a construct and measurement of privacy interest. We integrated the first-order dimensions of student interest, grounded in the cognitive model of empowerment, and theorized the stages of development of privacy interest from situational interest triggered by a privacy event to dispositional interest, whereby privacy protection becomes inherent in a consumer's behavior. Similar to student interest, we omitted the choice dimension from interest and instead included awareness to adapt interest to a privacy context. Privacy interest includes awareness, meaningfulness, competence, and impact as four first-order indicators to the higher-order privacy interest construct. We assessed the validity and reliability of our measurement scale following the scale development guidelines recommended by MacKenzie et al. (2011) and updated by Weber (2021). As a result, our study contributes to the information privacy research discourse by offering a complementary lens through which to view consumers' introspective thoughts of privacy protection (privacy interest) in addition to their reactions to companies and technologies (privacy concern), giving privacy researchers an expanded foundation on which to build privacy discourse.

2.7. Appendix 2A – Review of Articles Citing the Extended Privacy Calculus

We conducted our literature review by carefully reviewing articles from the Senior Scholars' List of Premier Journals (*DSS*, *EJIS*, *I&M*, *I&O*, *ISJ*, *ISR*, *JAIS*, *JIT*, *JMIS*, *JSIS*, *MISQ*) that cited the extended privacy calculus model for e-commerce transactions (Dinev & Hart, 2006). Formally, Dinev and Hart (2006) theorized and empirically tested *personal Internet interest* as a predictor to behavioral intentions in a privacy calculus decision-making context.

Table A1. Summary of Articles Included in Our Literature Review on the Extended Privacy Calculus

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Adjerid et al. (2018b) MISQ	✓					Protection satisfaction, harm perception, privacy notice	Hypothetical disclosure, actual disclosure, hypothetical + actual disclosure
Alashoor et al. (2022) ISR	✓					Cognitive resource, mood state	Disclosure behaviors
Al-Natour et al. (2021) JMIS		✓	✓	✓	✓	Design characteristics (why explanations, how explanations, speech acts), perceived transparency, perceived responsiveness, perceived interdependence (covariation of interests)	Intentions to self-disclose
Aloysius et al. (2013) ISR						Pricing	Seller profit, consumer value
Anderson and Agarwal (2011) ISR	✓	✓			✓	Risk scenario variables (type of information, intended purpose, requesting stakeholder) Health status emotion	Willingness to provide access to personal health info
Angst and Agarwal (2009) MISQ	✓					Argument frame, issue involvement, ability, post-attitude, pre-attitude	Opt-in intention
Bansal and Nah (2022) I&M	✓				✓	Right to be forgotten, oversight from surveillance	Trust propensity
Bansal et al. (2010) DSS	✓	✓			✓	Poor health status, perceived health information sensitivity, Big 5 personality, previous online privacy invasion, past positive experience with the website	Intention to disclose health information

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Bansal et al. (2015) EJIS	✓				✓	Argument quality in ELM (adequacy: collection; errors; secondary use; improper access), peripheral cues in ELM (availability of company info, website info quality, design appeal, reputation)	Intention to disclose privacy information
Bansal et al. (2016) I&M	✓				✓	Previous online privacy invasion, Big 5 personality, prior positive experience with the website	Intention to disclose information
Bélanger and Crossler (2011) MISQ	✓					Group dynamics, group information privacy concern, societal information privacy concern, individual differences, government involvement, organization information privacy concern, organizational environment	Four types of information privacy concerns (individual, group, organization, societal)
Belanger and Crossler (2019) JSIS	✓				✓	Mobile protection settings awareness, prior invasion experience, attitude towards information sharing, mobile privacy protection self-efficacy, mobile information protection intention	Mobile information protection (behavior)
Bélanger and James (2020) ISR	✓	✓		✓		Time, environmental characteristics, salient social identity, information privacy norms, information privacy norm development, multilevel information privacy decision, experiential feedback	Multilevel information privacy decision and behavior
Benlian et al. (2020) ISJ						Intrusive technology features (unintentional voice activation, presenteeism, anonymity), anthropomorphic technology features (anthropomorphic design), privacy invasion	Strain, interpersonal conflict
Breward et al. (2017) ISR	✓			✓	✓	Familiarity, perceived control, account security, convenience (benefit), security concerns	Attitude
Buckman et al. (2019) ISR						Gender, age, education, false information, web usage, breach history, information context, secondary use, identifying information	Willingness to accept (privacy valuation)
Cavusoglu et al. (2016) ISR						Time frame (short run, long run), policy change, gender, friendship network	Disclosure behaviors (wall posts, private messages)
Cheikh-Ammar (2020) I&M	✓	✓				Escape (from reality, to fantasy), SNS well-being (competence, relatedness, autonomy), SNS enjoyment, SNS overload (excessive demand, invasion), SNS exhaustion	Intention
Chen (2013) DSS	✓	✓				Social presence, ease of use, extroversion, internet risk perception, enjoyment	Site use
Chen et al. (2021b) I&M	✓	✓			✓	Disposition to value privacy, perceived cyber attack, exposure, attitude, age, internet use, length of SNS use, usefulness of SNS	Site use

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Cheng et al. (2021) I&M		✓		✓		Privacy awareness, previous privacy invasion, mobile payment security, negative media exposure, personal information disclosure requirements, immediate gratification, intention to disclose information	Disclosure of personal information
Choi and Land (2016) I&M	✓					Information collection, profile control	Willingness to delegate profile to Facebook apps
Choi et al. (2015) ISR						Perceived privacy invasion, perceived relationship bonding, information dissemination, network commonality	Inaction, avoidance (transactional, interpersonal), approach
Choi et al. (2016) JMIS						Justice perceptions (procedural, distributive, interaction), perceived breach, feelings of violation, pre-incident outcomes (word-of-mouth, likelihood of switching)	Post word-of-mouth, post-likelihood of switching
Choi et al. (2018) JAIS	✓	✓				Network mutuality, profile diagnosticity, expected social capital gains	No-action, acceptance
Cichy et al. (2021) MISQ	✓	✓			✓	Data sensitivity, data security, psychological ownership, self-efficacy enhancement, self-image congruency	Sharing of personal driving data
Conger et al. (2013) ISJ	✓	✓		✓	✓	First party (consumer/individual transaction: product/service, consumer, environment, medium, vendor, information, and social context characteristics), second party (vendor/provider of products and services), third party (legal data haring partners), fourth party (illegal entities)	Decision calculus
Crossler and Posey (2017) JAIS	✓				✓	Interpersonal characteristics (censorship attitude, self-efficacy, behavioral-based inertia, previous similar experience), reputation, perceived system characteristics (system granularity, system efficacy, perceived inconvenience), web activity, web location, network type	Intentions to use identity ecosystem
Datta and Chatterjee (2008) EJIS					✓	Electronic market inefficiencies (anonymity, lack of product transparency, lack of process transparency) information specificity, uncertainty in electronic markets, agency costs	Need for institution-based trust in intermediaries
Davidson et al. (2018) I&O						Journal, year, health IT adoption and diffusion, physician resistance to health IT use, health IT impact on health care or system outcomes	Health care research publication in IS (2004 to 2018)
de Corbiere and Rowe (2013) JAIS						Structural linkages, shared data, flow of messages	Interconnections between sending and receiving systems
Dinev et al. (2008) JSIS	✓					Perceived need for government surveillance, government intrusion concerns	Willingness to provide personal information to transact on the internet

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Dinev et al. (2013) EJIS		✓		✓		Tactics of information control (anonymity, secrecy, confidentiality), information sensitivity, importance of information transparency, regulatory expectations, perceived information control	Perceived privacy
Dinev et al. (2015) ISR	✓	✓		✓	✓	Antecedents (privacy experiences, awareness, personality, demographics, differences, culture, climate) level of effort (affect, cognitive resources, motivation, time constraints), peripheral cues, biases, heuristics, misattribution	Behavioral reactions
Fernando Libaque-Saenz et al. (2021) I&M		✓				Fair information practices, automatic data collection, perceived data control,	Behavioral intention
Furneaux and Wade (2017) JMIS						Institutional norms, system capability shortcomings, system support availability, system investment, replacement risk, system complexity	Replacement intention
Galbreth and Shor (2010) MISQ						Malicious agents, quality differentiation, horizontal differentiation, market share, market coverage	Enterprise system adoption
Gerlach et al. (2015) JSIS		✓				Privacy policy permissiveness	Willingness to disclose
Gerlach et al. (2019) JAIS		✓				Stereotypical thinking about providers' handling of user information, response to a provider's privacy statement, misjudgment of a provider's user-information-handling activities	Privacy risk perceptions
Gu et al. (2017) DSS	✓					Perceived app popularity, perceived permission sensitivity, permission justification, mobile privacy victim experience	Download intention
Herath et al. (2014) ISJ	✓					Email risk perception, email screening self-efficacy, eAuth attitude (eAuth usefulness, eAuth ease of use, eAuth responsiveness), eAuth privacy notification practice	Coping motivation
Hoehle et al. (2015) EJIS						Mobile application usability (application design, application utility, interface graphics, interface structure, interface input, interface output), individual/collectivism, masculinity/femininity, power distance, long term orientation, uncertainty avoidance	Continued intention to use
Hoehle et al. (2019) EJIS	✓					Artifact design (hardware, content), mobile application usability (application content, user interaction, interface presentation)	Shopping efficiency (product evaluation cost, product screening cost, decision-making quality)

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Hong and Thong (2013) MISQ	✓	✓			✓	Interaction management (collection, secondary usage, control), information management (errors, improper access), awareness	Internet privacy concern
Hu et al. (2015) EJIS		✓		✓		Enjoyment, curiosity fulfillment, effort, online social value, satisfaction	Continued use
Huang et al. (2021) ISR			✓			Ex ante registration request, ex-post registration request, total number of purchases, total user revenue, screening of low-interest users	Registration, short-term conversion, long-term purchase behavior
Hui et al. (2007) MISQ	✓				✓	Privacy statements, privacy seals, monetary incentives, information request, internet shopping experience, information misuse experience, cookie preference setting	Information disclosure
James et al. (2015) I&M						Interpersonal privacy identity (information management, interaction management), privacy calculus (information seeking, socialization, self-expression, pleasing others)	Information and interaction management behaviors (contact, profile, work, introspective, extrospective)
James et al. (2017) I&M	✓					Individualistic/collectivistic cultural orientation, Facebook information disclosure self-efficacy, severity of exposing others, susceptibility of others to exposure	Use of Facebook privacy controls
Jia et al. (2022) MISQ			✓			Social communication, attention to detail, structure, technicity, gender, age, employment, education	Personal innovativeness in IT (intrinsic interest in IT)
Jiang et al. (2013) ISR	✓			✓		Perceived anonymity of self, perceived anonymity of others, perceived media richness, perceived intrusiveness, social rewards (benefit)	Self-disclosure, misrepresentation
Junglas et al. (2008) EJIS	✓					Big 5 personality	Concern for privacy
Karwatzki et al. (2017a) JMIS						Personalization, disposition to value privacy, transparency features	Intention to disclose
Karwatzki et al. (2017b) EJIS						Physical, social, resource-related, psychological, prosecution-related, career-related, freedom-related	Perceived adverse consequences of access to individuals' information
Karwatzki et al. (2022) ISJ		✓		✓		Disposition to value privacy, privacy experience	Willingness to provide information
Kehr et al. (2015) ISJ	✓	✓		✓	✓	Affect, information sensitivity, perceived privacy	Intention to disclose
Keith et al. (2015) ISJ	✓	✓		✓	✓	Task/action coping efforts, structural assurances, disposition to trust, mobile-computing self-efficacy	Actual disclosure

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Koh et al. (2020) DSS				✓		Costs of disclosing information (opt-out from the firm's email solicitation list, opt-out from the firm's text message solicitation list), brand loyalty (number of visits to the website, number of purchases), information disclosure, demographic information	Coupon redemption
Koohikamali et al. (2015) DSS	✓	✓		✓		Social norm, opinion leadership, attitude toward location based social network applications, incentives, facilitating conditions	Location disclosure on location based social network applications
Kordzadeh and Warren (2017) JAIS	✓			✓		Expected positive personal outcomes of communicating PHI, expected positive community outcomes of communicating PHI, affective commitment, gender, age	Willingness to communicate personal health information
Krasnova et al. (2010) JIT		✓		✓	✓	Perceived control, convenience, relationship building, self-presentation, enjoyment (benefit)	Self-disclosure
Kummer et al. (2018) DSS		✓				Extroversion, conditional value, receivers, frequency, location relevance, disclosure value	Disclosure intention
Kwak et al. (2019) JAIS						SNS addiction, perceived threat severity, perceived threat susceptibility	Digital piracy intention, perceived usefulness, perceived ease of use
Lee et al. (2011a) MISQ						Consumer reservation value, consumers' preference intensity parameters, consumer location in terms of the preference, proportion of the unconcerned, proportion of pragmatists, proportion of fundamentalists, personalization scope parameter, price of standard product, privacy of personalized product, gathering cost, investment cost of privacy protection, fixed cost to protect privacy	Profit of firm
Lee et al. (2011b) DSS						Information security protection level, firm revenue, profit, profit-at-risk, overall implementation cost, loss severity, frequency of information security breaches, total financial losses, expected financial losses	Information security investment
Leidner and Tona (2021) MISQ						Dignity (behavioral, meritocratic, inherent), empowerment, privacy, emancipation, identity, personal data digitalization (knowing-self, showing-self, knowing-others, showing-others)	Responses to dignity disequilibrium (micro-level: forfeit, fight, flight, befriending, tending; macro-level: mobilize, comply, resist, regulate)

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Li (2012) DSS	✓	✓		✓	✓	Procedural fairness, social contract (trust), social response, social presence, information boundary, personalities, threat appraisal: perception of intrusion, coping appraisal, attitude toward disclosure, subjective norm for disclosure, perceived behavioral control: privacy self-efficacy	Intention to disclose, disclosure behavior
Li and Karahanna (2015) JAIS						Understand consumer (consumer information collection, building consumer profile), deliver recommendations (matchmaking approaches, recommendation system presentation), personalized recommendation	Impacts of recommendation system
Li and Unger (2012) EJIS	✓			✓		Privacy protection, perceived quality of personalization (benefits), industry domain, past experience, likelihood of using online personalization	Willingness to pay a premium, willingness to provide info
Li et al. (2015) I&M						Demographics, social network site experience personal social network size, plogging productivity	Privacy disclosure behaviors
Li et al. (2017) I&M	✓					Motive consistency (environmental dimension), perceived privacy control (interpersonal dimension), liking (environmental dimension)	Behavioral intention
Lin and Armstrong (2019) JAIS	✓	✓			✓	Territory coordination (linkage, permeability, ownership), information sensitivity	Private disclosure
Lin et al. (2017) I&M		✓		✓		Confirmation, perceived usefulness, satisfaction, perceived enjoyment (benefits), perceived reputation, community identification, gender	SNS continuance intention
Lin et al. (2021) EJIS	✓	✓			✓	Relative advantage, perceived ease of use, compatibility	Intention to use, use
Liu and Wang (2018) I&M		✓		✓		Group norms, perceived effectiveness of privacy settings, role conflict, role overload, disposition to value privacy, social rewards (benefits), privacy control, US vs. China	Intention to self-disclose
Liu et al. (2016) I&M		✓		✓	✓	Perceived anonymity self, benefits (convenience of relationship maintenance, relationship building, enjoyment, self-presentation)	Self-disclosure (amount, depth, honest, intent, valence)
Liu et al. (2019) ISJ		✓				Habit, role conflict, emotion, perceived control	Self-disclosure
Liu et al. (2020) I&M	✓					Role conflict, role overload, social interaction anxiety, disappointment	Lurking intention

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Liu et al. (2022) ISR	✓				✓	Privacy policy design (non-negotiation privacy policy application, negation, non-active-recommendation privacy policy application, negation, active-recommendation privacy policy application), demographics, years of using banking app, banking app usage, reputation, mobile privacy experience	Disclosure intention, disclosure behavior
Lowry et al. (2011) JMIS	✓					Masculinity, uncertainty avoidance, power distance, collectivism, desire for awareness, privacy victim, gender, age, education, attitude toward IM technology, behavioral intention to use IM	Use of instant messaging
Lowry et al. (2013) JMIS		✓			✓	The failure ought to be reported, responsibility to report the failure, confidence in WBRS anonymity	Willingness to report the failure
Mai et al. (2010) JMIS						Reputation, awareness, competition, attention, number of items, click-and-mortar presence, social technology, market competition, vendor reviews, customer reviews, critics reviews, product selection, best seller's list, loyalty rewards, subscriptions, personalized recommendations, birthday recommendation, security seal	Price premium
Mettler and Wulf (2019) ISJ		✓		✓		Physiolytics scenarios, prototypes, system properties, affordances, constraints	User types of physiolytics at the workplace
Miltgen and Smith (2015) I&M		✓		✓	✓	Regulatory knowledge, perceived privacy regulatory protection, perceived rewards (benefits)	Protection behavior, regulatory preferences
Miltgen and Smith (2019) I&M		✓		✓	✓	Perceived relevance	Context-specific consequences: privacy protective behavior (withholding, falsification)
Mirzaei and Esmaeilzadeh (2021) I&M						experience with online health community features, experience with peers, experience with disease, experience with OHC culture, perceived channel richness, informational support, emotional support, willingness to share information, willingness to seek information, perceived health status, engagement in OHC	Self-care efficacy, health outcome
Moody et al. (2017) EJIS					✓	Distrust (malevolence, incompetence, deceit)	Overall intentions
Ogbanufe and Gerhart (2020) ISJ	✓			✓		Benefits (belongingness, social interactions), IT smartwatch identity	Deep use, innovative individual performance

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Ozdemir et al. (2017) EJIS	✓	✓		✓	✓	Privacy experiences, privacy awareness	Information disclosure
Parks et al. (2017) EJIS						Enacting privacy safeguards, evaluating privacy safeguards enactments (intended and unintended consequences), imbalance challenge, workarounds & reactance (ignoring encryption, borrowing password, unattended logged on computers)	Impact on privacy compliance
Pavlou (2011) MISQ	✓	✓		✓	✓	Information privacy practices, information privacy tools and technologies, levels of analysis, sample characteristics,	Information privacy concern
Raddatz et al. EJIS	✓			✓		Blockchain awareness, perceived threat severity of storing information on a nonblockchain database, perceived threat susceptibility of storing information on a nonblockchain databased, inertia in switching to blockchain (affective, behavioral, cognitive)	Intention to switch to blockchain
Schwaig et al. (2013) I&M	✓					Self-esteem, consumer alienation, computer anxiety, attitude (permission, transfer, technology)	Behavioral intention
Shen et al. (2019) I&M			✓			Technology attractiveness (task, social, physical), interest in social commerce (interest), community involvement	Social commerce engagement
Sheng et al. (2008) JAIS	✓					Personalization, context	Intention to adopt
Shih et al. (2017) EJIS					✓	Cognitive social identity, affective social identity, evaluative social identity, switching cost, dependency	Online self-disclosure
Smith et al. (2011) MISQ	✓	✓		✓	✓	Privacy experiences, privacy awareness, personality differences, demographic differences, culture/climate, regulation, privacy notice/seal	Behavioral reactions (including disclosures)
Son and Kim (2008) MISQ	✓			✓		Perceived justice (interactional, procedural, distributive)	Information provision (refusal, misrepresentation), privacy action (removal, negative word-of-mouth), public action (complaining directly to online companies, complaining indirectly to third-party organizations)
Spiekermann and Korunovska (2017) JIT		✓				Market awareness (asset consciousness, market awareness x org. privacy), technical market design (data use control), engagement, psychological ownership (efficacy, identity, home, friends), market morality (privacy accessibility risk), technical market design (data storage redundancy)	Monetary valuation of personal data and personal data appreciation

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Sun et al. (2021b) I&M		✓		✓		Dual motivation system (behavioral activation system, behavioral inhibition system), reward responsiveness BAS drive, BAS fun seeking, life documentation, self-expression, social rewards	Disclosure intention
Tang and Ning (2023) DSS	✓			✓		Perceived privacy control, disposition to value privacy, perceived app permission sensitivity, perceived effectiveness of privacy policies, social rewards	Misrepresentation behavior
Teubner and Flath (2019) JAIS	✓			✓		Log(Audience size), perceived audience size, personal connection	Intention to share
Trenz et al. (2018) I&M						Positive WOM, negative WOM, peer use, subjective norm, uncertainty, service diagnosticity, internet experience, gender, age	Continued use intention
Tsai et al. (2011) ISR	✓					Price with shipping, privacy level, privacy icon, non-privacy sensitive item, privacy-sensitive items	Purchase price
Turel and Qahri-Saremi (2023) ISJ							Attitudinal response (ambivalence avoidance, negative response amplification, positive response amplification, concessions, holism)
Vance et al. (2014) JAIS						Willingness to gamble lifetime income, general risk appetite, perceived security risk of malware, threat susceptibility, threat severity, bias, malware warning screen realism, hacker screen realism, malware warning screen concern, hacker screen concern, demographics	Security warning disregard
Wakefield (2013) JSIS	✓				✓	Internet security, positive affect (enjoy), negative affect	Intentions to disclose
Wall et al. (2016) JAIS		✓				Formal and information communication structures, violation coupling, enforceability (certainty, severity, celerity of sanctions), goal clarity of rules, rule connectedness, economic and noneconomic strain	Likelihood of a privacy or security rule violation
Wang and Wu (2014) I&M						ESP proactive privacy governance (proactive provision and protection, proactive education, proactive monitor & feedback seeking), perceived value (emotional, social, functional), value-added strategies	Disclosure willingness for U-services
Wang et al. (2016) I&M					✓	Information quality, system quality, service quality, perceived value, customer satisfaction, relationship commitment	Stickiness intention

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Warkentin et al. (2017) JAIS	✓	✓		✓	✓	Psychological ownership, social influence, meter invasiveness, program discount (benefit), third party access	Behavioral intention
Wattal et al. (2012) ISR						Product-based personalization, personalized greeting, familiarity, promotion characteristics, prior response, prior purchase, frequency of e-mail, prior response	Probability of opening an email, response
Wottrich et al. (2018) DSS	✓					App intrusiveness, perceived app value	Permission acceptance intention
Wright et al. (2014) ISR						Liking, reciprocity, social proof, consistency, authority, scarcity, fictitious shared experience, self-determination, gender	Phishing response
Wu and Luo (2022) I&M						Privacy seal, acquisition time, shipping time, in stock, public, brand, channel, traffic, online age, uninformedness, product assortment	Likelihood of charging maximum and minimum list prices
Xu et al. (2009) JMIS		✓		✓		Information delivery mechanisms, compensation, industry self-regulation, government regulation	Intention to disclose personal information in location-based service
Xu et al. (2011a) JAIS	✓	✓				Institutional privacy assurance (perceived effectiveness of privacy policy, perceived effectiveness of industry self-regulation), disposition to value privacy, privacy control	Privacy concerns
Xu et al. (2011b) DSS		✓		✓		Willingness to have personal information used in LAM, previous privacy experience, personal innovativeness, coupon proneness, perceived value of info disclosure, covert vs. overt, personalization	Purchase intention
Xu et al. (2012b) ISR	✓					Personal agency control (individual self-protection), proxy control agency (industry self-regulation, government legislation), perceived control over personal information	Context-specific concerns for information privacy
Xu et al. (2015) DSS		✓		✓		App utility, app quality, aesthetics, enjoyment, knowledge of alternative quality, technicality, non-monetary sacrifices, satisfaction, perceived price, app continuance intention, intention to recommend	Recommendation
Yaraghi et al. (2019) JAIS		✓		✓		Number of patient medical records, number of patient medical providers, stigmatized medical conditions, medical provider tenure, location	Consent

Citation (journal)	PC	PR	IN	BN	TR	Other constructs	DV
Yun et al. (2019) I&M	✓	✓			✓	Demographics, internet literacy / experience, self-efficacy, personality, perceived vulnerability, privacy regulation, perceived control, disposition to privacy, privacy invasion, information sensitivity, privacy awareness, cultural values	Intention to provide personal information, intention to transact online, privacy setting / management, negative behaviors (avoidance), attitude toward technology, usefulness
Zhang et al. (2018) I&M	✓			✓		Threat appraisals (perceived severity, perceived vulnerability), coping appraisal (response efficacy, self-efficacy), perceived health status	PHI disclosure intention
Zhang et al. (2022) MISQ	✓			✓		Perceived intrusiveness, social rewards (benefits), age, gender, Facebook experience, privacy violation experience, misrepresentation of identity	Self-disclosure
Zhu et al. (2017) I&M		✓		✓		Store credits, free gift / trial samples, product discounts, vouchers, accuracy of recommendation, spam mails / crank calls, risk of privacy being traded, risk of account being hacked, time consumed, discriminatory pricing	Pricing strategy
Zhu et al. (2023) JAIS		✓		✓		Control, contexts	State of privacy
Zimmer et al. (2010a) DSS	✓			✓	✓	Dyadic condition, intention	Actual behavior
Zimmer et al. (2010b) I&M		✓			✓	Relevance, attitude, usefulness	Intent to disclose, actual disclosure (not investigated)

Note. PC = privacy concerns; PR = privacy risk perceptions; IN = interest; BN = benefits; TR = trust perceptions. The "interest" constructs identified in the studies pertained to a general interest toward technology, website, digital services, and the like. None of the articles included a conceptualization or operationalization of a privacy interest construct.

Journal names: DSS = Decision Support Systems; EJIS = European Journal of Information Systems; I&M = Information & Management; I&O = Information and Organization; ISJ = Information Systems Journal; ISR = Information Systems Research; JAIS = Journal of the Association for Information Systems; JIT = Journal of Information Technology; JMIS = Journal of Management Information Systems; JSIS = Journal of Strategic Information Systems; MISQ = MIS Quarterly

2.8. Appendix 2B – Survey and Measurement Details

Table B1. Wave 1 Instrument Independent and Control Variables

Variable	Contextualized items	Original items	Source
Age screen	Please indicate your age in years: **(Must be 18 or above)**		**Adapted from Hoehle and Venkatesh (2015)**

Prolific ID	Please provide your Prolific ID:		
Privacy interest (Choice)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your choices around protecting your information privacy**: 1. I typically choose to protect **my information privacy**. 2. I have a choice in the methods I can use to safeguard **my information privacy**. 3. I have a choice when it comes to defending **my information privacy**. 4. I have freedom to choose how to protect **my information privacy.** 5. The extent to which my **information privacy is protected** depends on my choices. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Choice 1. I have a choice in the methods I can use to perform my work. 2. I have freedom to choose among options in this class. 3. I can determine how tasks can be performed. 4. I have no freedom to choose in this class. 5. Alternative approaches to learning are encouraged in this class. 6. I have the opportunity to contribute to the learning of others in this class. (Items are presented in their original form to provide context of how privacy items were adapted based on the collection of items and not on any individual item)	Adapted from Frymier et al. (1996) *student interest* scale, which was derived from Schultz and Shulman (1993) *learner empowerment* scale based on Thomas and Velthouse (1990) conceptualization of *empowerment*
Privacy interest (Impact)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding the impact you can have on safeguarding your information privacy**: 1. My participation is impactful to the success of **safeguarding my information privacy**. 2. I can make an impact on what happens to **my information privacy**. 3. My contribution to **protecting my information privacy** makes a difference. 4. I can make an impact on the way **my information privacy** is safeguarded. 5. I make a difference in **protecting my information privacy**. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Impact 1. My participation is important to the success of this class. 2. I cannot influence what happens in this class. 3. My contribution to this class makes no difference. 4. I can make an impact on the way things are run in this class. 5. I make a difference in the learning that goes on in this class. 6. I have the power to make a difference in how things are done in this class. 7. I have the opportunity to make important decisions in this class. 8. I have the power to create a supportive learning environment in this class. 9. I can influence the instructor. 10. I feel appreciated in this class. (Items are presented in their original form to provide context of how privacy items were adapted based on the collection of items and not on any individual item)	Same as above

Privacy interest (Meaningfulness)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **on how meaningful your information privacy is to you**: 1. The task required of me to **protect my information privacy** is personally meaningful. 2. I find **information privacy protection** to be meaningful to me. 3. Information privacy protection is important to me. 4. Information privacy protection is personally meaningful. 5. **My information privacy** is very important to me. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Meaningfulness 1. The tasks required of me in this class are personally meaningful. 2. I look forward to going to this class. 3. This class is exciting. 4. This class is boring. 5. This class is interesting. 6. The tasks required of me in this class are valuable to me. 7. The information in this class is useful. 8. This course will help me achieve my future goals. 9. The tasks required in this course are a waste of my time. 10. This class is not important to me. (Items are presented in their original form to provide context of how privacy items were adapted based on the collection of items and not on any individual item)	Same as above
Privacy interest (Competence)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your competency in safeguarding your information privacy**: " I . . ." 1. ". . . am able to perform the necessary actions to **protect my information privacy**." 2. ". . . feel confident when it comes to my ability to **safeguard my information privacy**." 3. ". . . possess the necessary skills to successfully **protect my information privacy**." 4. ". . . believe that I am capable of **safeguarding my information privacy**." 5. ". . . know the steps I need to take to **protect my information privacy**." [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Competence 1. I feel confident that I can adequately perform my duties. 2. I feel intimidated by what is required of me in this class. 3. I possess the necessary skills to perform successfully in class. 4. I feel unable to do the work in this class. 5. I believe that I am capable of achieving my goals in this class. 6. I have faith in my ability to do well in this class. 7. I have the qualifications to succeed in this class. 8. I lack confidence in my ability to perform the tasks in this class. 9. I feel very competent in this class. (Items are presented in their original form to provide context of how privacy items were adapted based on the collection of items and not on any individual item)	Same as above
Internet privacy	**Prompt**: "Read carefully and indicate your agreement	1. It usually bothers me when	Hong and Thong

concern (Collection)	with each of the following statements **regarding the collection of your personal information**: 1. It usually bothers me when **websites** ask me for personal information. 2. When **websites** ask me for personal information, I sometimes think twice before providing it. 3. I am concerned that **websites** are collecting too much personal information about me. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	**commercial/government websites** ask me for personal information. 2. When **commercial/government websites** ask me for personal information, I sometimes think twice before providing it. 3. I am concerned that **commercial/government websites** are collecting too much personal information about me. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	(2013)
Internet privacy concern (Secondary Usage)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding the use of your personal information**: 1. I am concerned that when I give personal information to a **website** for some reason, the website would use the information for other reasons. 2. I am concerned that **websites** would sell my personal information in their computer database to other companies. 3. I am concerned that **websites** would share my personal information with other companies without my authorization. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	1. I am concerned that when I give personal information to a **commercial/government website** for some reason, the website would use the information for other reasons. 2. I am concerned that **commercial/government websites** would sell my personal information in their computer databases to other companies. 3. I am concerned that **commercial/government websites** would share my personal information with other companies without my authorization. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Hong and Thong (2013)
Internet privacy concern (Errors)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **on errors found in your personal information**: 1. I am concerned that **websites** do not take enough steps to make sure that my personal information in their files is accurate. 2. I am concerned that **websites** do not have adequate procedures to correct errors in my personal information.	1. I am concerned that **commercial/government websites** do not take enough steps to make sure that my personal information in their files is accurate. 2. I am concerned that **commercial/government websites** do not have adequate procedures to correct errors in my personal information. 3. I am concerned that **commercial/government websites** do not devote enough time and effort to verifying the accuracy of my personal information in their databases.	Hong and Thong (2013)

	3. I am concerned that **websites** do not devote enough time and effort to verifying the accuracy of my personal information in their databases. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	[Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	
Internet privacy concern (Improper Access)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding access to your personal information**: 1. I am concerned that databases that contain my personal information are not protected from unauthorized access. 2. I am concerned that **websites** do not devote enough time and effort to preventing unauthorized access to my personal information. 3. I am concerned that **websites** do not take enough steps to make sure that unauthorized people cannot access my personal information in their computers. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	1. I am concerned that **commercial/government website** databases that contain my personal information are not protected from unauthorized access. 2. I am concerned that **commercial/government websites** do not devote enough time and effort to preventing unauthorized access to my personal information. 3. I am concerned that **commercial/government websites** do not take enough steps to make sure that unauthorized people cannot access my personal information in their computers. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Hong and Thong (2013)
Internet privacy concern (Control)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding control over your personal information**: 1. It usually bothers me when I do not have control of the personal information that I provide to **websites**. 2. 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 **websites**. 3. I am concerned when control is lost or unwillingly reduced as a result of a marketing transaction with **websites**.	1. It usually bothers me when I do not have control of personal information that I provide to **commercial/government websites**. 2. 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 **commercial/government websites**. 3. I am concerned when control is lost or unwillingly reduced as a result of a marketing transaction with **commercial/government websites**. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Hong and Thong (2013)

	[Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]		
Internet privacy concern (Awareness)	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your awareness of data collection practices**: 1. I am concerned when a clear and conspicuous disclosure is not included in online privacy policies of **websites**. 2. It usually bothers me when I am not aware or knowledgeable about how my personal information will be used by **websites**. 3. It usually bothers me when **websites** seeking my information online do not disclose the way the data are collected, processed, and used. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	1. I am concerned when a clear and conspicuous disclosure is not included in online privacy policies of **commercial/government websites**. 2. It usually bothers me when I am not aware or knowledgeable about how my personal information will be used by **commercial/ government websites**. 3. It usually bothers me when **commercial/government websites** seeking my information online do not disclose the way the data are collected, processed, and used. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Hong et al. (2021); Hong and Thong (2013)
Global Information Privacy Concern	**Prompt**: "Read carefully and indicate your agreement with each of the following statements about **information privacy**: 1. Compared to others, I am more sensitive about the way online companies handle my personal information. 2. To me, it is the most important thing to keep my privacy intact from online companies. 3. I am concerned about threats to my personal privacy today. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	1. Compared to others, I am more sensitive about the way online companies handle my personal information. 2. To me, it is the most important thing to keep my privacy intact from online companies. 3. I am concerned about threats to my personal privacy today. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Malhotra et al. (2004)
Risk beliefs	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your risk perceptions toward using websites**: 1. In general, it would be risky to give my personal information to **websites**. 2. There would be high potential for loss associated with giving my personal information to **websites**.	1. In general, it would be risky to give my personal information to **commercial/government websites**. 2. There would be high potential for loss associated with giving my personal information to **commercial/government websites**. 3. There would be too much uncertainty associated with giving my personal information to **commercial/government websites**. 4. Providing **commercial/government websites** with	Hong and Thong (2013)

	3. There would be too much uncertainty associated with giving my personal information to **websites**. 4. Providing **websites** with my personal information would involve many unexpected problems. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	my personal information would involve many unexpected problems. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	
Trusting beliefs	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your trust perceptions toward websites**: 1. **Websites** in general would be trustworthy in handling my personal information. 2. **Websites** would keep my best interests in mind when dealing with my personal information. 3. **Websites** would fulfill their promises related to my personal information. 4. **Websites** are in general predictable and consistent regarding the usage of my personal information. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	1. **Commercial/Government websites** in general would be trustworthy in handling my personal information. 2. **Commercial/Government websites** would keep my best interests in mind when dealing with my personal information. 3. **Commercial/Government websites** would fulfill their promises related to my personal information. 4. **Commercial/Government websites** are in general predictable and consistent regarding the usage of my personal information. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Hong and Thong (2013)
Perceived benefit	**Prompt**: "Read carefully and indicate to which extent each of the following statements **describes your feelings** about **the benefits of using websites**: 1. I think using **websites** is convenient. 2. I can save money by using **websites**. 3. I can save time by using **websites**. 4. Using **websites** enables me to accomplish tasks more quickly. 5. Using **websites** increases my productivity in making decisions or finding information within the shortest time frame. [Likert-type 7-point scale: 1 = Does not describe my feelings: 7 = Completely describes my feelings]	1. I think using **this website** is convenient. 2. I can save money by using **this website**. 3. I can save time by using **this website**. 4. Using **this website** enables me to accomplish a shopping task more quickly than using traditional stores. 5. Using **this website** increases my productivity in shopping (e.g., making purchase decisions or finding product information within the shortest time frame). [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	Kim et al. (2008)

Blue attitude marker variable	**Prompt**: "Indicate your color preference. I . . ." 4. ". . . prefer blue to other colors" 5. ". . . like the color blue" 6. ". . . like blue clothes" [Likert-type 7-point scale: 1 = Does not describe my feelings: 7 = Completely describes my feelings]	1. I prefer blue to other colors 2. I like the color blue 3. I like blue clothes	Miller and Chiodo (2008)
Control variables			
Misrepresentation of identification	Same as original	Some websites ask for you to register with the site by providing personal information. When asked for such information, what percent of the time do you falsify the information? 1 = I have never falsified information 2 = under 25% of the time 3 = 26%–50% of the time 4 = 51%–75% of the time 5 = over 75% of the time	Malhotra et al. (2004)
Internet experience	Please indicate your Internet experience (in years):	1 = less than a year 2 = 1–less than 2 years 3 = 2–less than 3 years 4 = 3–less than 4 years 5 = 4–less than 5 years 6 = 5–less than 6 years 7 = 6–less than 7 years 8 = more than 7 years	Malhotra et al. (2004)
Prior privacy experience	Same as original	1. How often have you personally experienced incidents whereby your personal information was used by some company or e-commerce web site without your authorization? 2. How much have you heard or read during the last year about the use and potential misuse of the information collected from the Internet? 3. How often have you personally been the victim of what you felt was an improper invasion of privacy? [1 = not very often; 7 = very often]	Xu et al. (2012a)
Response set item	1. If fish cannot walk on land, select "Somewhat disagree". 2. Select "Somewhat agree" for this item. 3. If there is one "h" in "three" then select		

	"Strongly disagree". 4. If the Atlantic Ocean is the name of a real ocean, select "Neither agree nor disagree". [Random set of response items presented to respondent]		

Table B2. Wave 2 Survey Procedures and Instructions

Procedure	Content
Prolific ID	Please provide your Prolific ID:
Instruction	**Instructions:** On the next page, we will ask you questions about your decision-making when interacting with online companies, their websites, or their mobile applications. We appreciate your honest and transparent responses. Please click the *arrow* to proceed.

Table B3. Wave 2 Instrument Dependent Variables, Demographics, and Internet Experience

Variable	Contextualized items	Original items	Source
Attention check	Q1.1 We have a little test to identify people who are paying attention to the instructions of the survey. In the next page, you will see a photo of one person. Then, you will be asked to state how many people you see in the photo. We want you to answer "three" even though you will see one person (Elon Musk) in the photo. This is to **identify the Prolific respondents** who pay attention to these instructions. In other words, in order to continue to the rest of the study and be paid for your time, you must answer incorrectly by choosing the "three" option. All right? Ok, please proceed to the next page! Q1.2 How many people can you see in this picture? (0; 1; 2; 3; 4; 5) [If 3 Is Not Selected, Then Skip To End of Survey]	Q1.1 Before we **start the task**, we have a little test to identify people who are paying attention to the instructions of the survey. In the next page, you will see a photo of three people. Then, you will be asked to state how many people you see in the photo. We want you to answer "three" even though you will see four people in the photo. This is to make sure that only **M-Turkers** who pay attention to these instructions continue to our task. In other words, in order to continue to the rest of the study and be paid for your time, you must answer incorrectly by choosing the "three" option. All right? Ok, please proceed to the next page! Q1.2 How many people can you see in this picture? (0; 1; 2; 3; 4; 5) [If 3 Is Not Selected, Then Skip To End of Survey]	Acquisti and Fong (2019)
Self-disclosure	**Prompt**: "Read carefully and indicate your agreement with each of the following statements **regarding your disclosure behaviors on websites**: 1. I **reveal** a great amount of information about myself to **websites**.	1. In the particular experience, I **revealed** a great amount of information about myself to **the other party**. 2. In the particular experience, I **gave** out intimate information to **the other party**.	Jiang et al. (2013)

	2. I **give** out intimate information to **websites**. 3. I **share** a variety of information about myself to **websites**. 4. I **disclose** information openly to **websites**. 5. I **reveal** very personal thoughts, feelings and experiences to **websites**. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	3. In the particular experience, I **shared** a variety of information about myself to **the other party**. 4. In the particular experience, I **disclosed** information openly to **the other party**. 5. In the particular experience, I **revealed** very personal thoughts, feelings and experiences to **the other party**. [Likert-type 7-point scale: 1 = Strongly disagree; 7 = Strongly agree]	
Removal	Same as original	Please specify the extent to which you would take actions to have your information removed from online companies' database when your personal information was not properly handled. 1) Very unlikely/very likely 2) Not probable/probable 3) Impossible/possible [Seven-point semantic scale]	Son and Kim (2008)
Complaining Directly to Online Companies	Same as original	Please specify the extent to which you would write or call online companies to complain about the way they use personal information when your personal information was not properly handled. 1) Very unlikely/very likely 2) Not probable/probable 3) Impossible/possible [Seven-point semantic scale]	Son and Kim (2008)
Complaining Indirectly to Third-Party Organizations:	Same as original	Please specify the extent to which you would write or call an elected official or consumer organization to complain about the way online companies use personal information when your personal information was not properly handled. 1) Very unlikely/very likely 2) Not probable/probable 3) Impossible/possible [Seven-point semantic scale]	Son and Kim (2008)
Misrepresentation	Same as original	Please specify the extent to which you would falsify some of your personal information if it is asked for by online	Son and Kim (2008)

		companies within the next three years. 1. Very unlikely/very likely 2. Not probable/probable 3. Impossible/possible [Seven-point semantic scale]	
Blue attitude marker variable	**Prompt**: "Indicate your color preference. 1) I like the color blue. 2) Blue is a beautiful color. 3) I enjoy the color blue. 4) Blue is a pleasant color. [Likert-type 7-point scale: 1 = Very untrue for me; 7 = Very true for me]	1) I like the color blue. 2) Blue is a beautiful color. 3) I enjoy the color blue. 4) Blue is a pleasant color. [Likert-type 7-point scale: 1 = Very untrue for me; 7 = Very true for me]	Miller and Chiodo (2008); adapted by Schuetz et al. (2021)
Response set item	1. If "2 + 3 = 6" then select "agree"; otherwise, select "strongly agree". 2. If you provided a Prolific ID to this survey, then select "neither agree nor disagree". [Random set of response items presented to respondent]		
Global privacy interest item	**Prompt**: "Indicate your level of interest in information privacy. I see why I should be interested in my information privacy in today's digital age. [Slider scale: 1 = Very untrue for me; 100 = Very true for me]	1. I don't see why we should learn the details of topics such as respiration of photosynthesis.	(Gardner & Tamir, 1989)
Open comment on privacy interest	Are you interested in your information privacy in today's digital age? Why or why not? (min. 100 characters)		n/a
Demographics			
Gender	Please indicate your identified gender: [Male / Female / Prefer not to say / Other [please specify]]		Adapted from Hoehle and Venkatesh (2015)
Education	What is the highest level of education you have completed? 1 = Less than high school / secondary school		Adapted from Al-Natour et al. (2020)

	2 = High school / secondary school 3 = Some university, but have not completed a degree 4 = Associate degree 5 = Bachelor's degree 6 = Master's degree 7 = Doctorate / Ph.D. [Radio button for each selection]		
Annual income range	What is your approximate annual income range? 1 = < $30,000 2 = $30,001 – $75,000 3 = $75,001 – $150,000 4 = $150,001 – $300,000 5 = $300,001 – $500,000 6 = $500,001+ [Radio button for each selection]		Adapted from Hoehle and Venkatesh (2015)
Employment status	Please indicate your current employment status: 1 = Employed part-time 2 = Employed full-time 3 = Not employed 4 = Self-employed 5 = Student 6 = Retired 7 = Other [Radio button for each selection]		Adapted from Al-Natour et al. (2020)
Ethnicity	Please indicate the ethnic group you most identify with: 1 = American Indian or Alaskan Native 2 = Asian 3 = Black or African American 4 = Hispanic or Latino 5 = Middle Eastern or North African 6 = White or Caucasian 7 = Prefer not to say [Radio button for each selection]		Adapted from Crossler and Bélanger (2019)

2.9. Appendix 2C – Descriptive Statistics of Control Variables and Demographics

Table C1. Descriptive Statistics of Control Variables in the Nomological Model

Constructs	Min/Max	Mean	Standard deviation
Age	19.00/82.00	42.89	13.62
Gender	0.00/1.00	45.2	0.50
Internet experience	4.00/8.00	7.96	0.35
Misrepresentation	1.00/5.00	2.20	1.07

Table C2. Distribution of Race

Race	Frequency	Percentage
American Indian or Alaskan Native	4	0.9
Asian	30	7.1
Black or African American	20	4.7
Hispanic or Latino	27	6.4
Middle Eastern or North African	1	0.2
Native Hawaiian or Pacific Islander	1	0.2
White or Caucasian	329	77.8
Prefer not to say	11	2.6

Table C3. Distribution of Employment

Employment	Frequency	Percentage
Employed part-time	52	12.3
Employed full-time	170	40.2
Not employed	71	16.8
Self-employed	66	15.6
Student	12	2.8
Retired	35	8.3
Other	17	4.0

Table C4. Distribution of Education

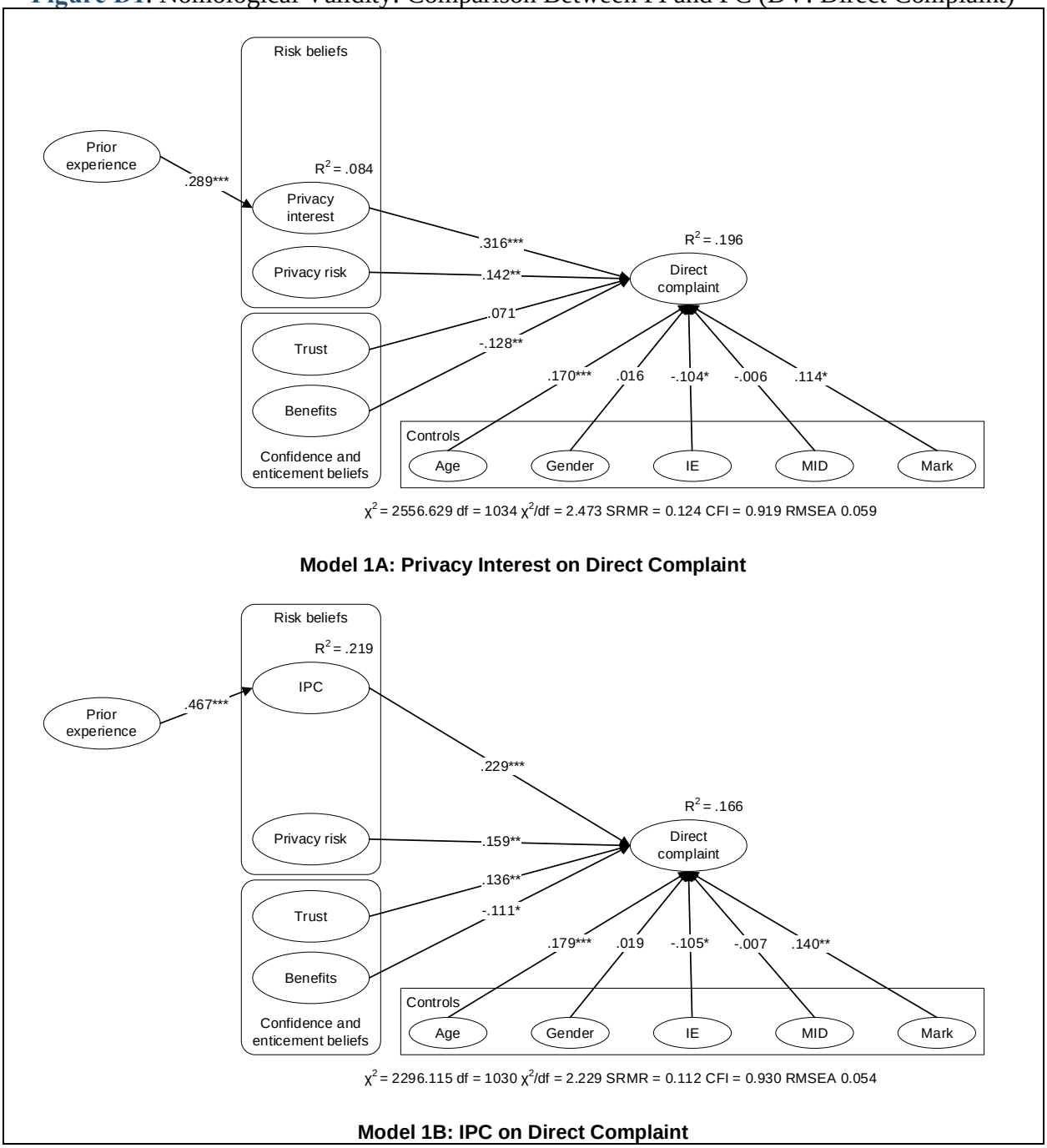
Education	Frequency	Percentage
Less than high school / secondary school	5	1.2
High school / secondary school	60	14.2
Some university, but have not completed a degree	90	21.3
Associate degree	55	13.0
Bachelor's degree	156	36.9
Master's degree	39	9.2
Doctorate / Ph.D.	18	4.3

Table C5. Distribution of Income

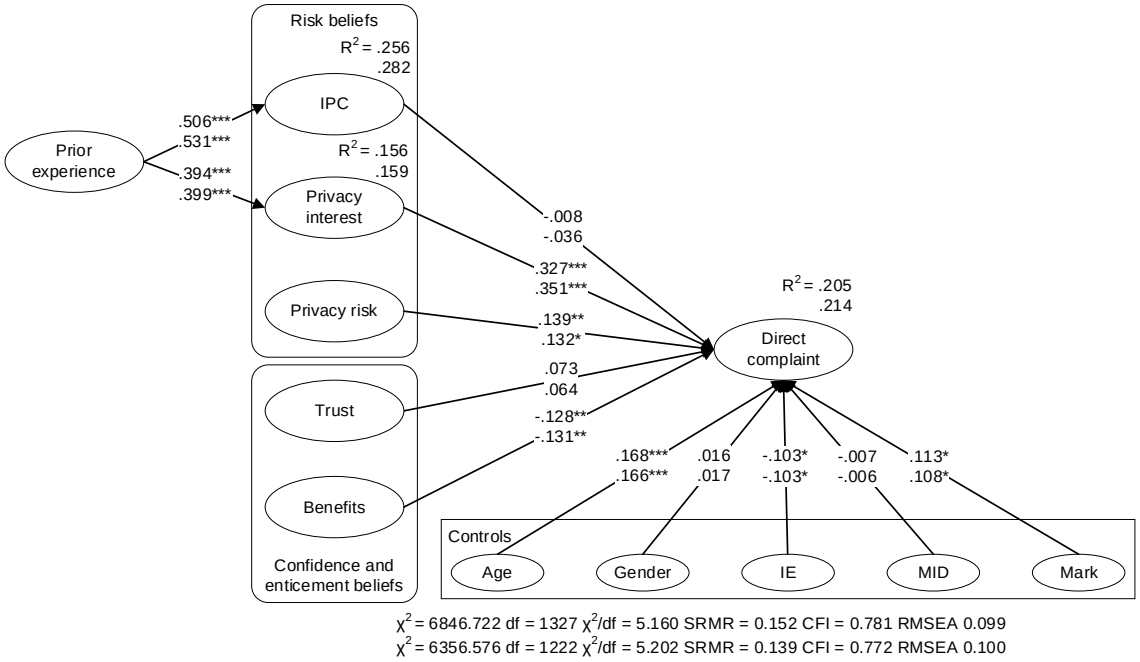
Income	Frequency	Percentage
< $30,000	170	40.2
$30,001 – $75,000	152	35.9
$75,001 – $150,000	83	19.6
$150,001 – $300,000	16	3.8
$300,001 – $500,000	2	0.5

2.10. Appendix 2D – Descriptive and Psychometric Properties

Figure D1. Nomological Validity: Comparison Between PI and PC (DV: Direct Complaint)

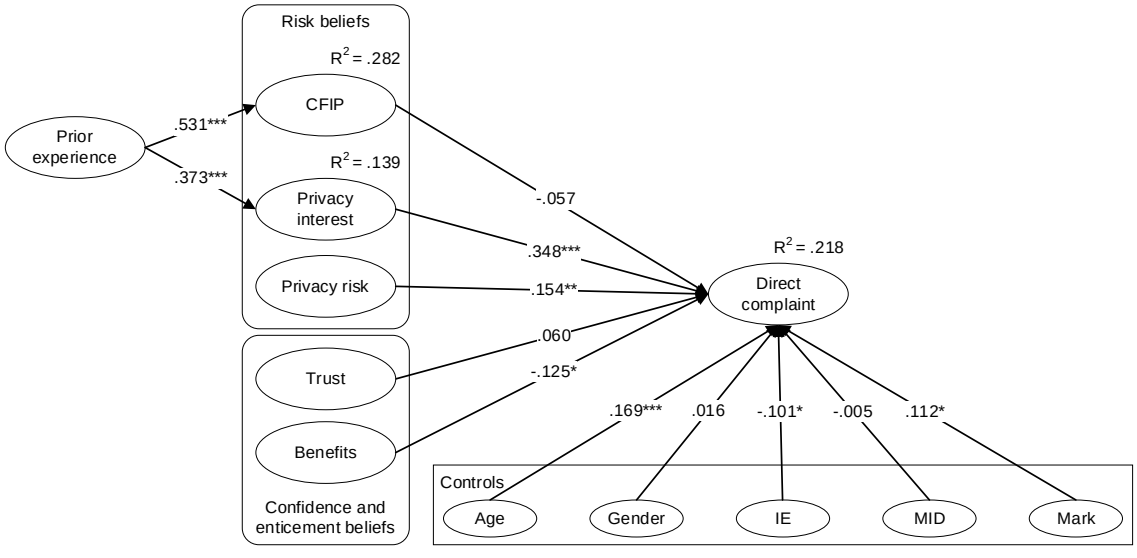


Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable



Model 1C & D: Privacy Interest and IPC on Direct Complaint

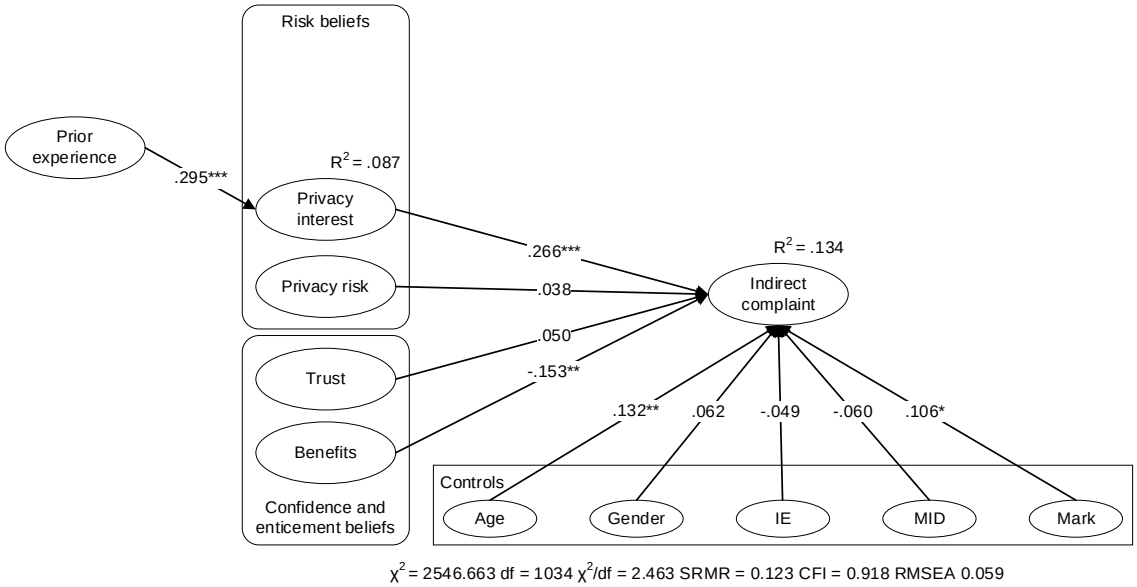
Upper-most listed values = Model C with IPC as *imputed factor score* variable and privacy interest as *latent* variable. Bottom-most listed values = Model D with IPC as *latent* variable and privacy interest as *imputed factor score* variable.



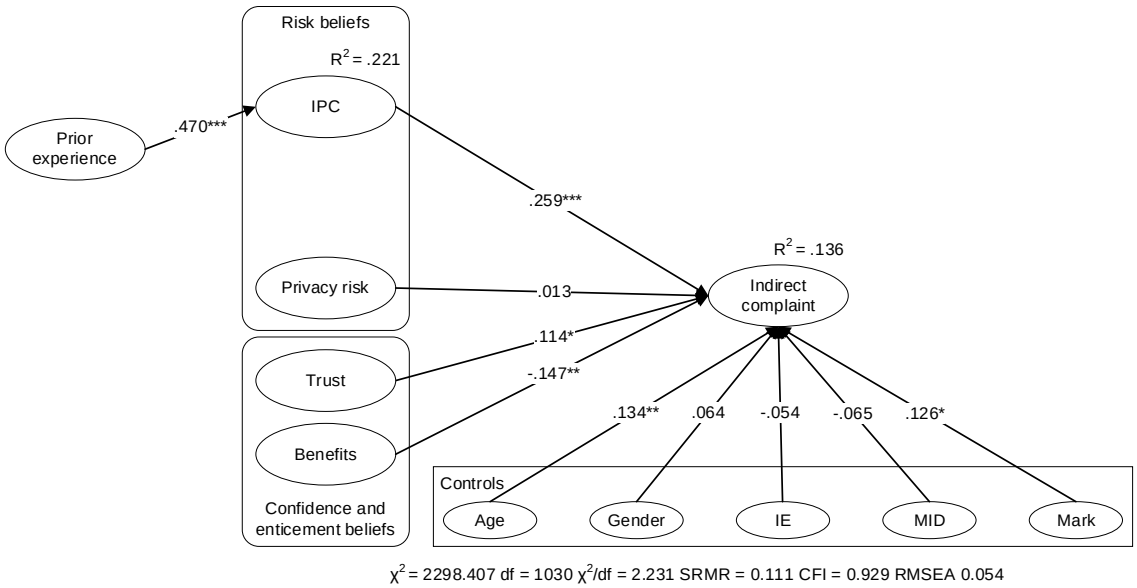
Model 1E: Privacy Interest and CFIP on Direct Complaint

Note. *p < 0.05, **p < 0.01, ***p < 0.001. CFIP = Concern for information privacy, IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable

Figure D2. Nomological Validity: Comparison Between PI and PC (DV: Indirect Complaint)

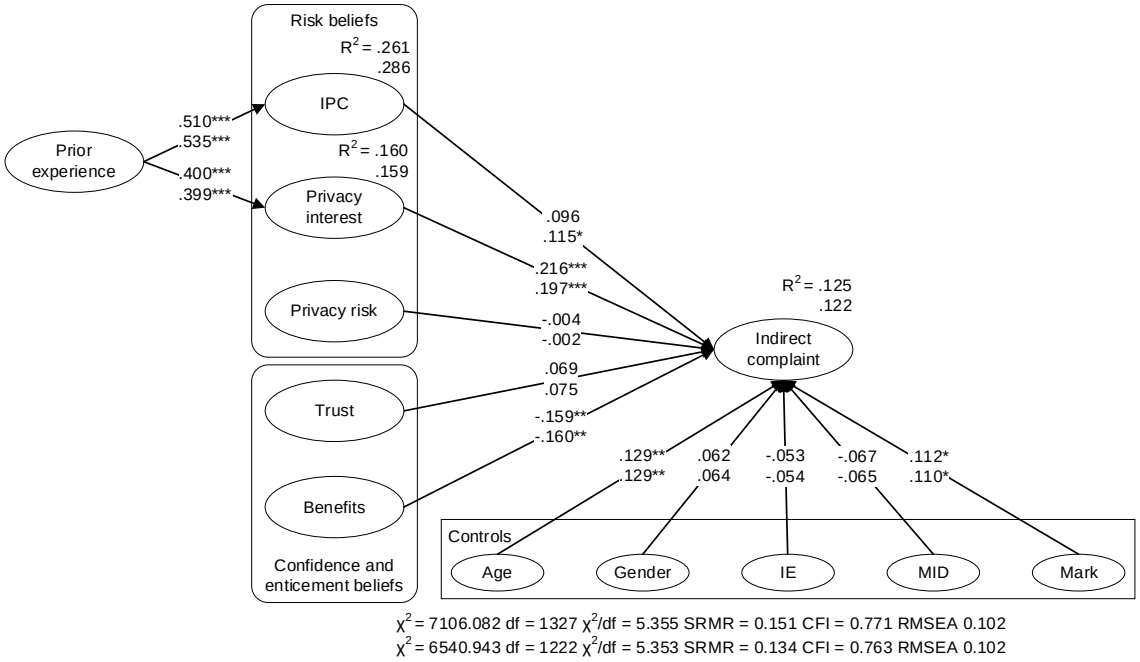


Model 2A: Privacy Interest on Indirect Complaint



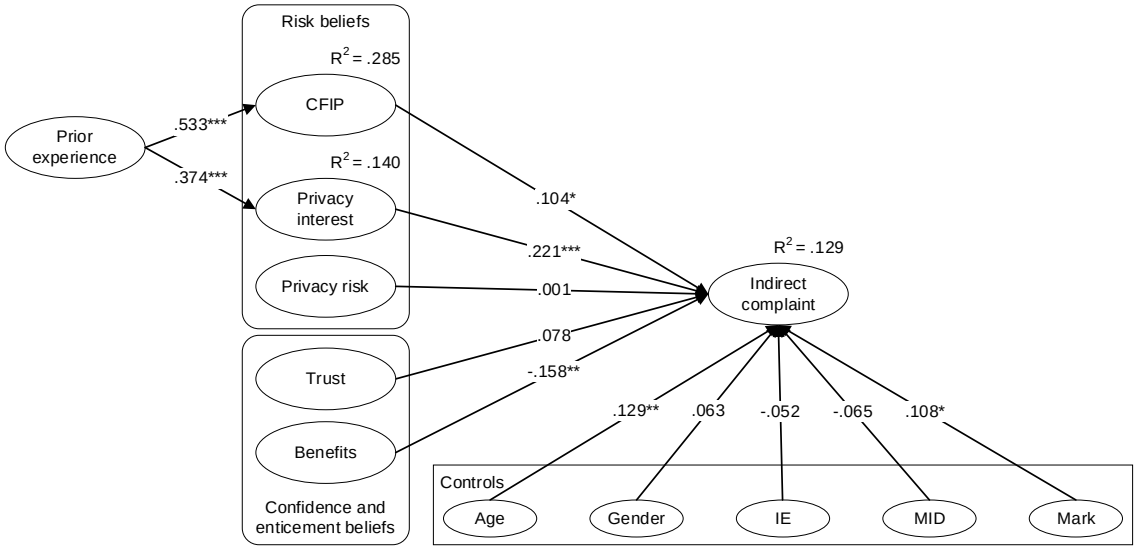
Model 2B: IPC on Indirect Complaint

Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable



Model 2C & D: Privacy Interest and IPC on Indirect Complaint

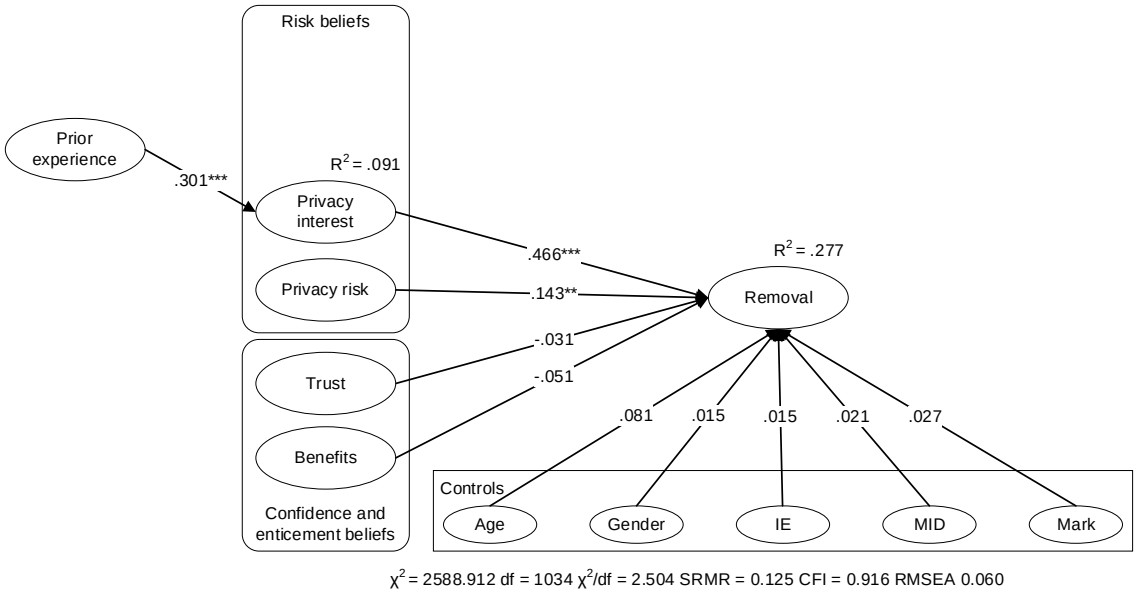
Upper-most listed values = Model C with IPC as imputed factor score variable and privacy interest as latent variable. Bottom-most listed values = Model D with IPC as latent variable and privacy interest as imputed factor score variable.



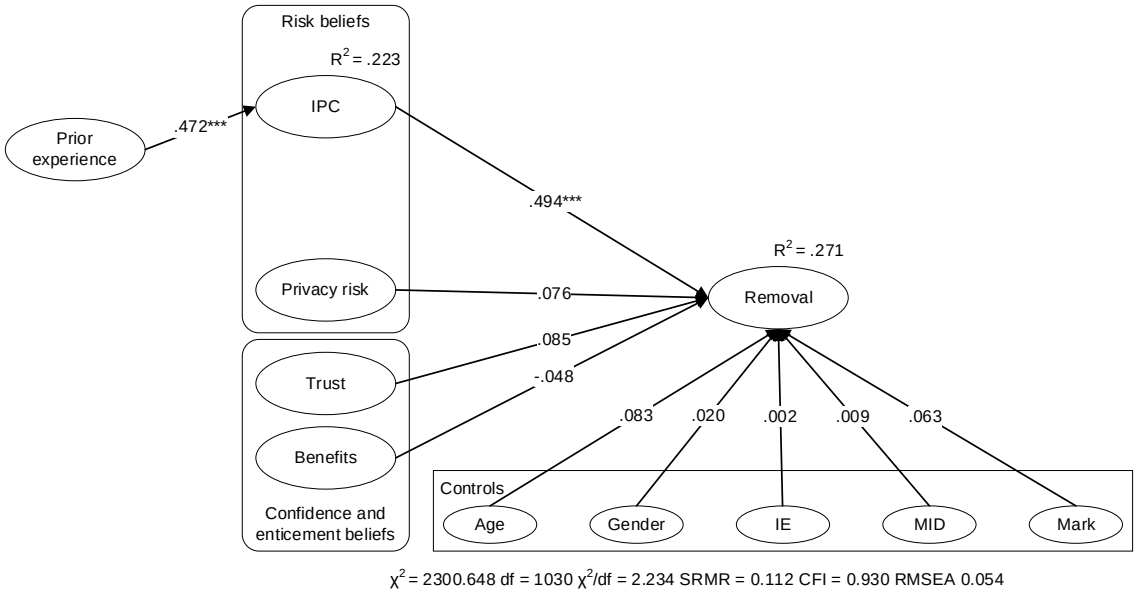
Model 2E: Privacy Interest and CFIP on Indirect Complaint

Note. *p < 0.05, **p < 0.01, ***p < 0.001. CFIP = Concern for information privacy, IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable

Figure D3. Nomological Validity: Comparison Between PI and PC (DV: Removal)

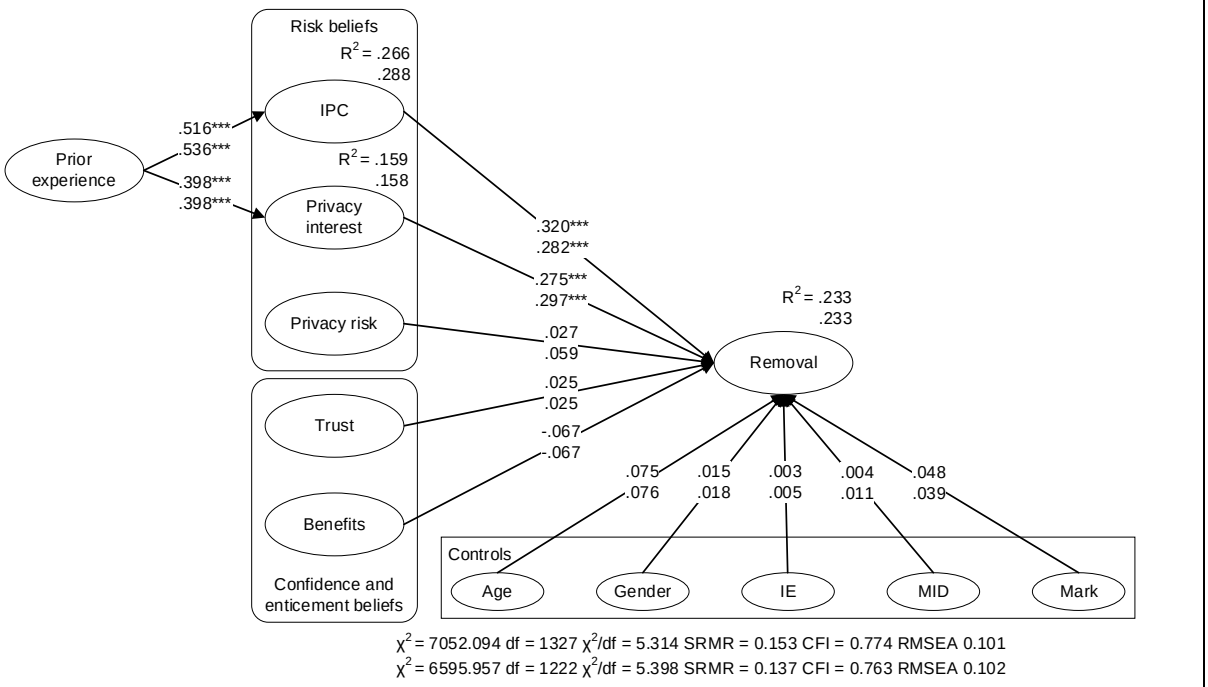


Model 3A: Privacy Interest on Removal



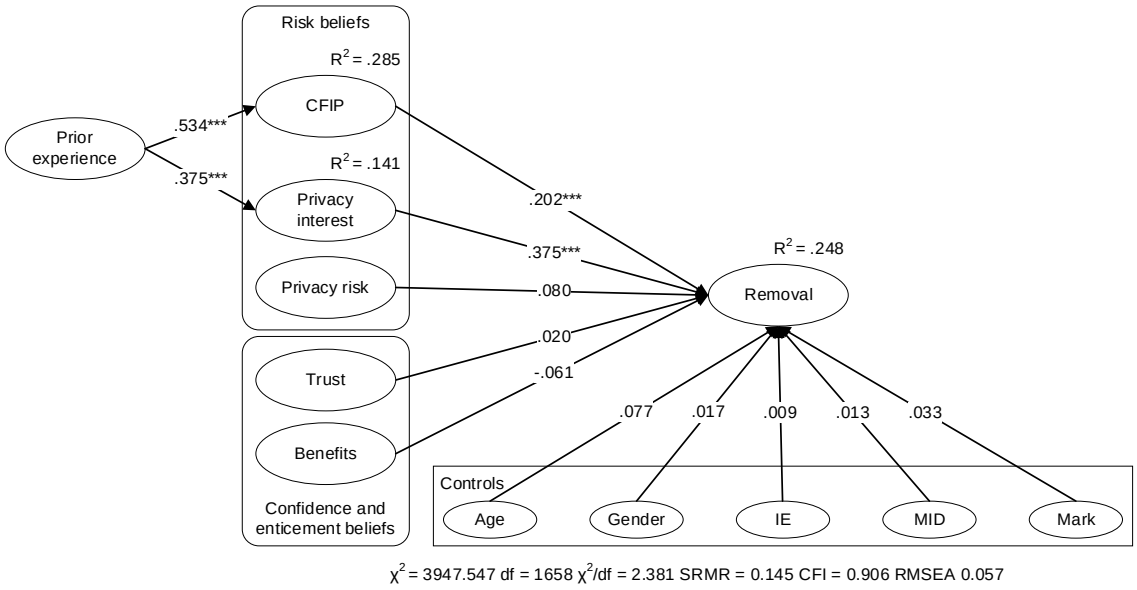
Model 3B: IPC on Removal

Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable



Model 3C & D: Privacy Interest and IPC on Removal

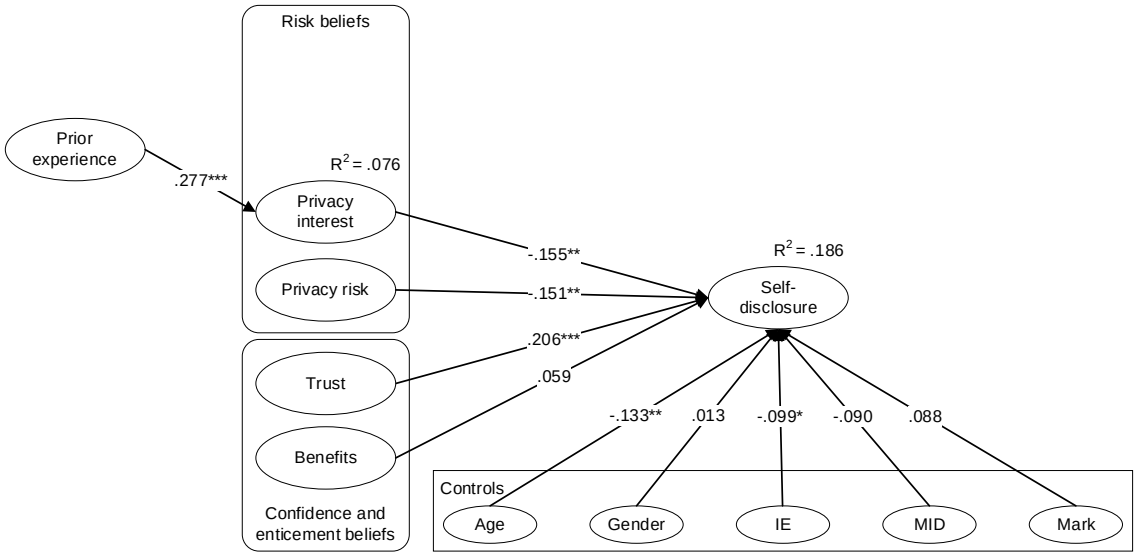
Upper-most listed values = Model C with IPC as *imputed factor score* variable and privacy interest as *latent* variable. Bottom-most listed values = Model D with IPC as *latent* variable and privacy interest as *imputed factor score* variable.



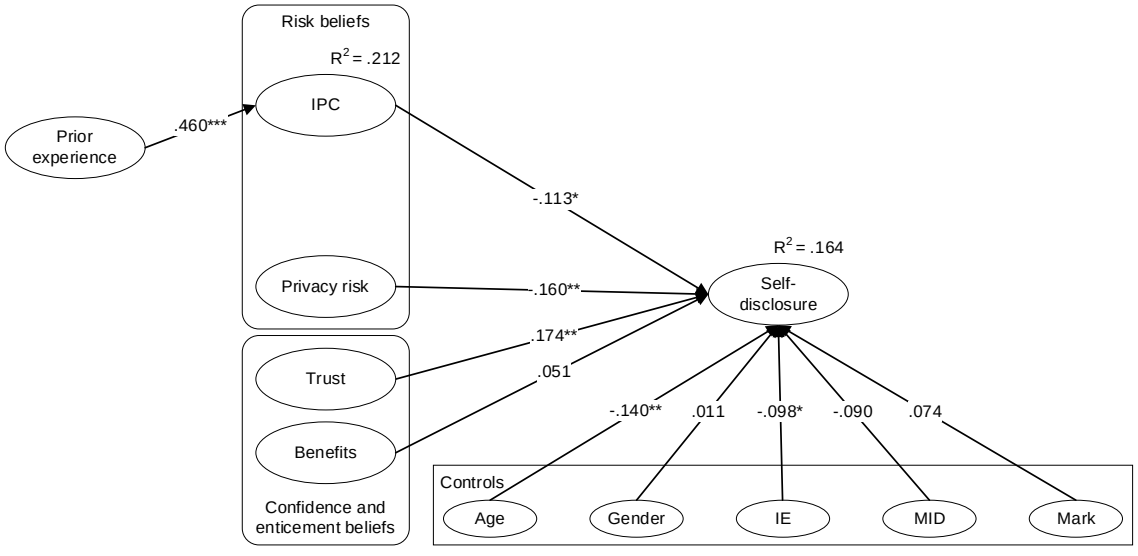
Model 3E: Privacy Interest and CFIP on Removal

Note. *p < 0.05, **p < 0.01, ***p < 0.001. CFIP = Concern for information privacy, IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable

Figure D4. Nomological Validity: Comparison Between PI and PC (DV: Self-disclosure)

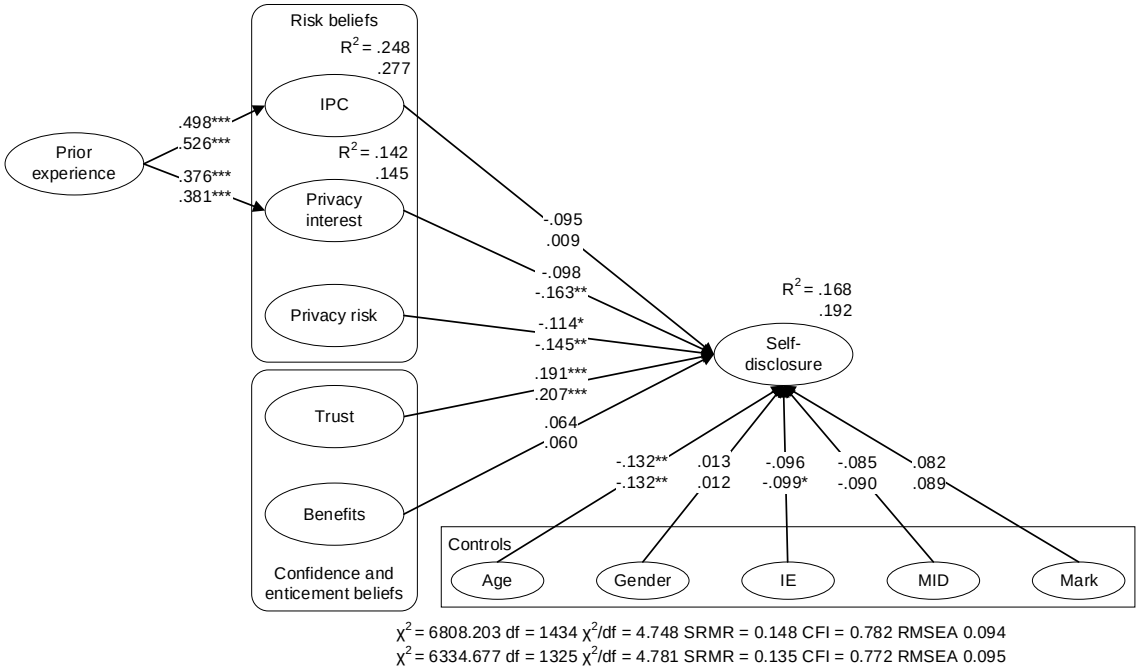


Model 4A: Privacy Interest on Self-disclosure



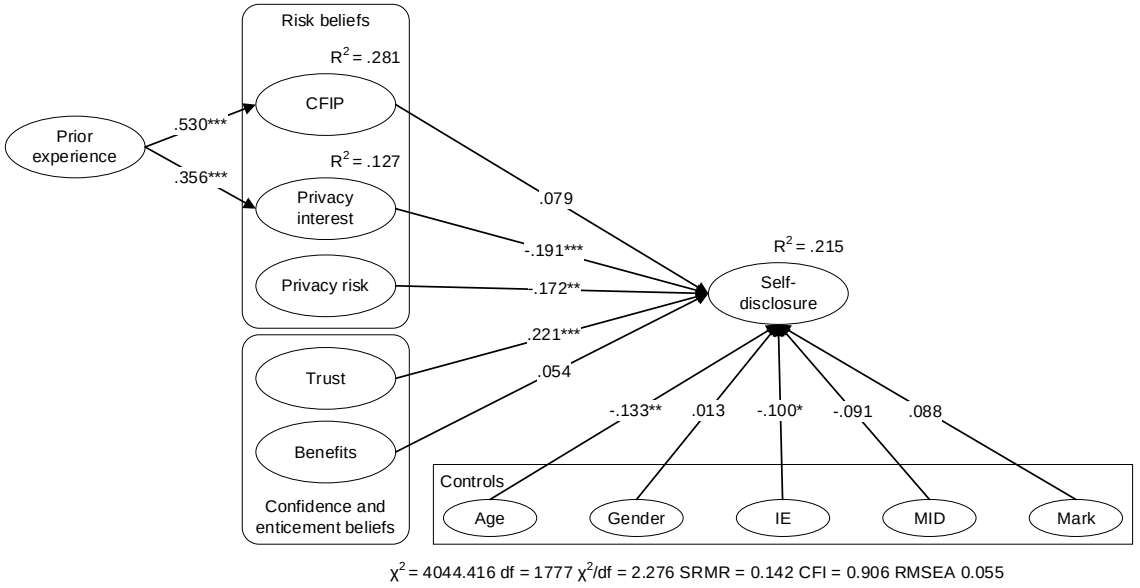
Model 4B: IPC on Self-disclosure

Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable



Model 4C & D: Privacy Interest and IPC on Self-disclosure

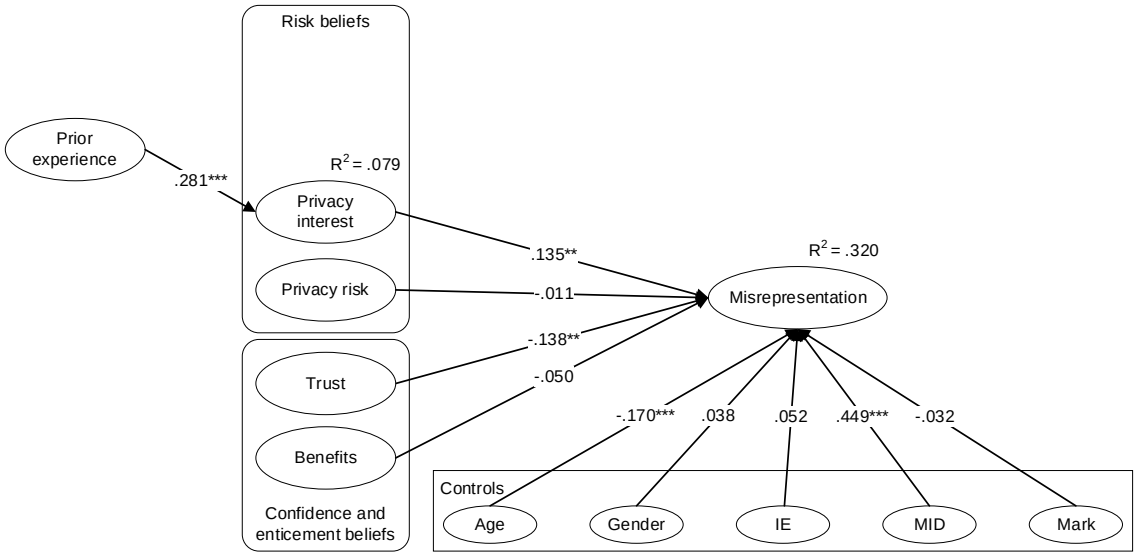
Upper-most listed values = Model C with IPC as *imputed factor score* variable and privacy interest as *latent* variable. Bottom-most listed values = Model D with IPC as *latent* variable and privacy interest as *imputed factor score* variable.



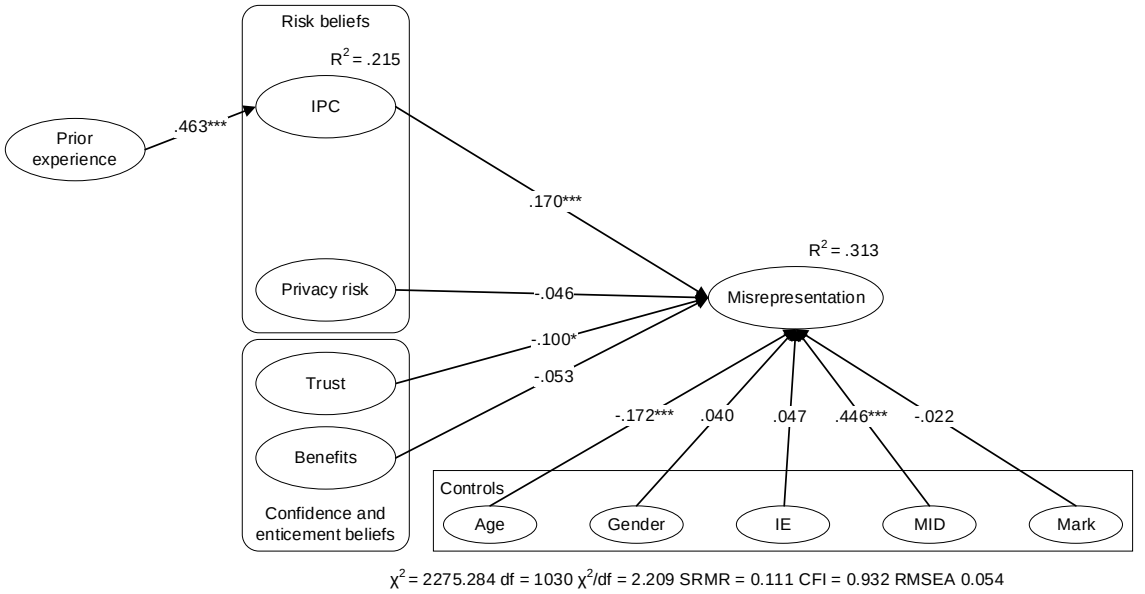
Model 4E: Privacy Interest and CFIP on Self-disclosure

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. CFIP = Concern for information privacy, IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable

Figure D5. Nomological Validity: Comparison Between PI and PC (DV: Misrepresentation)

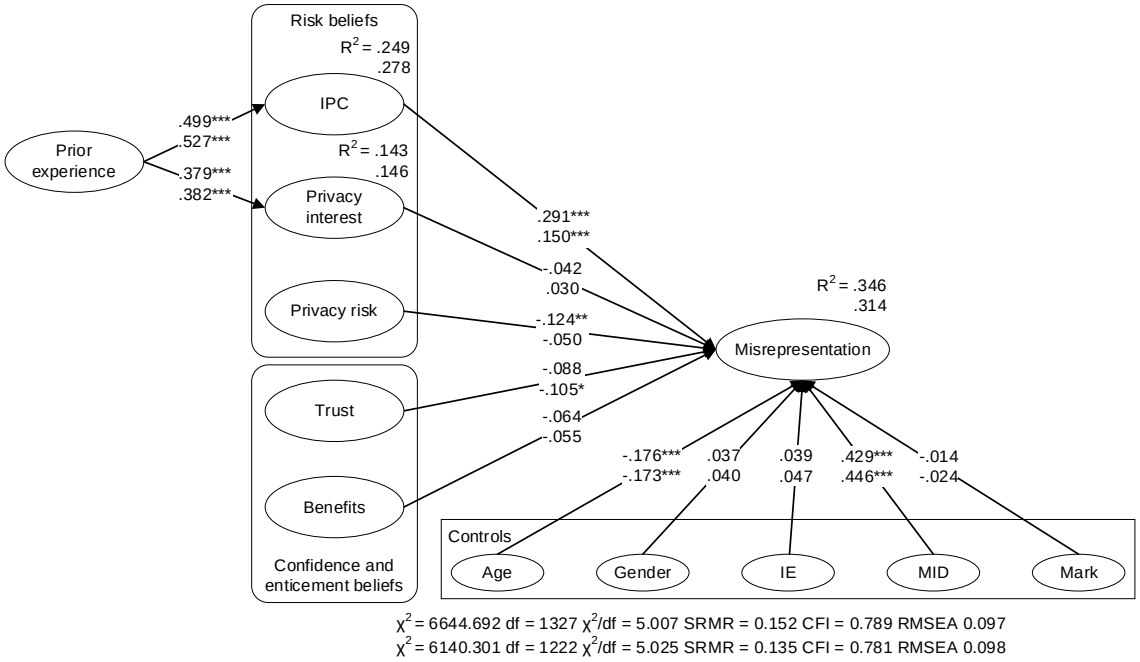


Model 5A: Privacy Interest on Misrepresentation



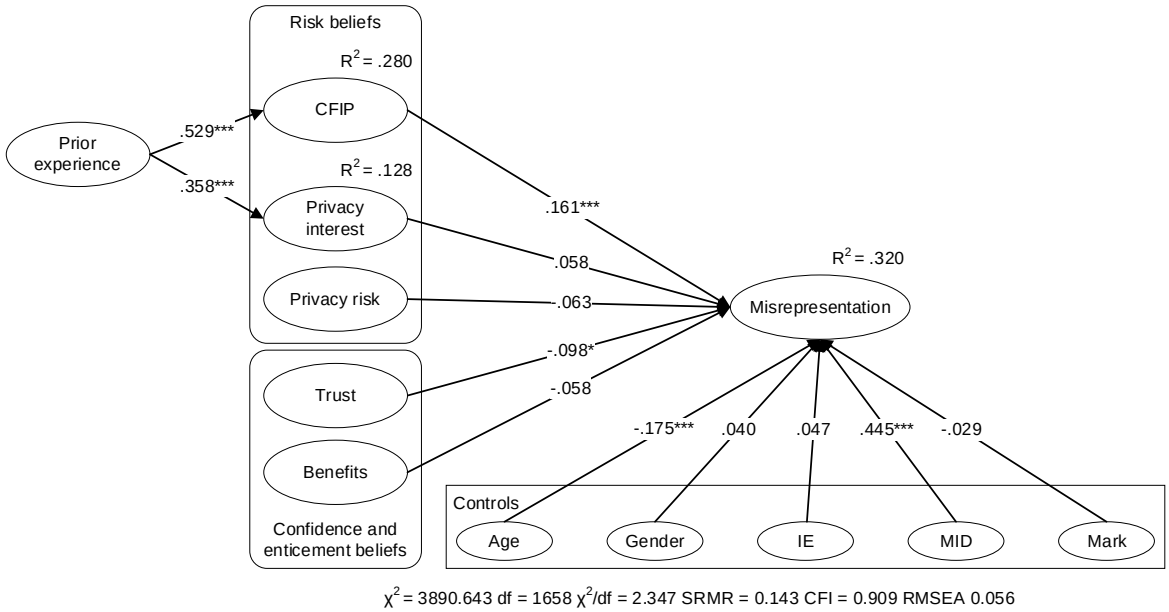
Model 5B: IPC on Misrepresentation

Note. *p < 0.05, **p < 0.01, ***p < 0.001. IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable



Model 5C & D: Privacy Interest and IPC on Misrepresentation

Upper-most listed values = Model C with IPC as *imputed factor score* variable and privacy interest as *latent* variable. Bottom-most listed values = Model D with IPC as *latent* variable and privacy interest as *imputed factor score* variable.



Model 5E: Privacy Interest and CFIP on Misrepresentation

Note. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$. CFIP = Concern for information privacy, IPC = Internet privacy concerns, IE = Internet experience, MID = Misrepresentation of identity, Mark = Marker variable

2.11. Appendix 2E – Content Validity Assessment

Two important coefficients to compute for content validity include (1) proportion of substantive agreement (P_{sa}) and (2) substantive validity coefficient (C_{sv}). The definitions for each are as follows:

1) ***Proportion of substantive agreement*** (P_{sa}): Indicates the proportion of respondents who assign items to their intended constructs (Anderson & Gerbing, 1991). For example, this is the number of people who assign a "Choice" item to the category "Choice." The formula is as follows:

$$P_{sa} = nc / N$$

"Where *nc* is the number of respondents who assigned an item to its intended construct. *N* is the total number of respondents. P_{sa} values range between 0 and 1. High values indicate that the construct definition represents the items judged" (Anderson & Gerbing, 1991).

2) ***Substantive validity coefficient*** (C_{sv}): Indicates the extent to which respondents assign items to the posited construct rather than to any other construct (Anderson & Gerbing, 1991). For example, "Choice_1" item may be grouped to "Choice," and "Impact." C_{sv} computes P_{sa} and subtracts the maximum number of times "Choice_1" was assigned to an incorrect construct. The formula is as follows:

$$C_{sv} = (nc - n0) / N$$

"Where nc is the number of respondents assigning an item to the intended construct, n0 is the highest number of assignment of the measure to any other construct [incorrect assignment], and N is the total number of respondents. C_{sv} values can range from -1 and 1. Positive values suggest that an item was assigned to its intended construct more than assignment to any other construct. Negative values suggest the opposite" (Hoehle & Venkatesh, 2015, p. 453).

The cut-off value for P_{sa} and C_{sv} is **0.60** (Hoehle & Venkatesh, 2015), which "suggests that **60%** of all raters associated the items with the intended construct definition" (p. 453). Table E1 shows the survey instrument.

Table E1. Content Validity Survey Instrument

Variable	Item	Source
Information sheet	"Information Sheet for Participation in a Research Study" requesting informed consent	
Age screen	Please indicate your age in years: (Must be 18 or above)	Adapted from Hoehle and Venkatesh (2015)
Instructions	A set of **40 privacy-related items** are shown below in the column to the left. Your objective is to carefully read each item. **Four concept groups** are located in the column to the right. A definition is provided for each concept. Please read each definition carefully. Each group represents a specific concept: *choice*, *competence*, *impact*, or *meaningfulness*. You will match the privacy-related item to the concept definition you feel most appropriately matches the item. Each item can belong to only one group, so select the most appropriate match between an individual item and its concept group. You will then be asked to explain why you matched those items to their respective concept group. *Each explanation requires a 200-character minimum response, so be mindful to document why you are placing the items into their respective groups.*	Anderson and Gerbing (1991); MacKenzie et al. (2011)
Item-matching activities	"List of 40 items for *Choice* (9 each), *Competence* (9 each), *Impact* (11 each), and *Meaningfulness* (11 each)" Respondents matched the items to one of four category groups: 1) **Choice**: Degree to which one feels' responsible for their privacy actions. 2) **Competence**: Degree to which one feels capable to perform the necessary activities to protect their information privacy. 3) **Impact**: Degree to which one's privacy behavior is seen as making a difference. 4) **Meaningfulness**: The value of performing privacy behaviors in relation to one's ideals, beliefs, or standards.	Anderson and Gerbing (1991); MacKenzie et al. (2011)
Response quality check	In your own words, please provide your reason for grouping the items into "**Choice**" (minimum 200 characters | maximum 1000 characters). [Open-text response field]	N/A
Response quality check	In your own words, please provide your reason for grouping the items into "**Competence**" (minimum 200 characters | maximum 1000 characters). [Open-text response field]	N/A
Response quality check	In your own words, please provide your reason for grouping the items into "**Impact**" (minimum 200 characters | maximum 1000 characters). [Open-text response field]	N/A
Response quality check	In your own words, please provide your reason for grouping the items into "**Meaningfulness**" (minimum 200 characters | maximum 1000 characters). [Open-text response field]	N/A

2.12. Appendix 2F – Common Latent Factor Assessment

Table F1. Item Loadings on the Common Latent Factor

Construct	Items	Loadings	
Awareness	AWA1	0.015	
	AWA2	0.043	
	AWA3	0.044	
Meaningfulness	MNG1	-0.088	
	MNG2	0.009	
	MNG3	0.043	
	MNG4	0.021	
	MNG5	0.054	
Impact	IMP1	0.067	
	IMP2	0.013	
	IMP3	0.026	
	IMP4	0.046	
	IMP5	0.010	
Competence	CMP1	0.009	
	CMP2	-0.152	
	CMP3	-0.018	
	CMP4	-0.020	
	CMP5	-0.017	
Collection	COL1	0.017	
	COL2	0.203	
	COL3	0.139	
Secondary Usage	USE1	0.129	
	USE2	0.136	
	USE3	0.145	
Errors	ERR1	-0.083	AVE = 0.012
	ERR2	-0.073	
	ERR3	-0.114	
Improper Access	ACC1	0.075	
	ACC2	0.074	
	ACC3	0.058	
Control	CTL1	0.041	
	CTL2	0.060	
	CTL3	0.106	
Risk Beliefs	RB1	-0.031	
	RB2	-0.092	
	RB3	-0.105	
	RB4	-0.124	
Trust Beliefs	TR1	-0.174	
	TR2	-0.219	
	TR3	-0.036	
	TR4	-0.097	
Perceived Benefits	BN1	0.192	
	BN2	0.016	
	BN3	0.101	
	BN4	0.098	
	BN5	0.129	
Self-disclosure	SD1	-0.012	
	SD2	-0.033	
	SD3	0.068	

	SD4	0.035	
	SD5	-0.038	
Removal from Company Database	REM1	-0.036	
	REM2	-0.034	
	REM3	0.289	
Direct Complaint to Company	COM1	-0.117	
	COM2	-0.083	
	COM3	0.297	
Indirect Complaint to Third-Party Organization	IND1	-0.067	
	IND2	-0.033	
	IND3	0.338	
Misrepresentation	MIS1	-0.009	
	MIS2	0.034	
	MIS3	0.259	
Prior Experience	PRI1	-0.021	
	PRI2	0.050	
	PRI3	0.003	
Blue Marker Variable	MRK1	-0.088	
	MRK2	0.054	
	MRK3	0.039	
	MRK4	0.090	
	MRK5	0.059	
	MRK6	0.062	
	MRK7	0.098	

Bibliography

Abramson, L. Y., Seligman, M. E., & Teasdale, J. D. (1978). Learned helplessness in humans: Critique and reformulation. *Journal of Abnormal Psychology, 87*, 49-74. https://doi.org/10.1037/0021-843X.87.1.49

Acquisti, A. (2004). *Privacy in electronic commerce and the economics of immediate gratification* Proceedings of the 5th ACM conference on Electronic commerce, New York, NY, USA.

Acquisti, A. (2009). Nudging privacy: The behavioral economics of personal information. *IEEE Security & Privacy, 7*(6), 82-85. https://doi.org/10.1109/MSP.2009.163

Acquisti, A., Adjerid, I., & Brandimarte, L. (2013). Gone in 15 seconds: The limits of privacy transparency and control. *IEEE Security & Privacy, 11*(4), 72-74. https://doi.org/10.1109/MSP.2013.86

Acquisti, A., Brandimarte, L., & Hancock, J. (2022). How privacy's past may shape its future. *Science, 375*(6578), 270-272. https://doi.org/10.1126/science.abj0826

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509-514. https://doi.org/10.1126/science.aaa1465

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: The drive for privacy and the difficulty of achieving it in the digital age. *Journal of Consumer Psychology, 30*(4), 736-758. https://doi.org/10.1002/jcpy.1191

Acquisti, A., & Fong, C. (2019). An experiment in hiring discrimination via online social networks. *Management Science, 66*(3), 1005-1024. https://doi.org/10.1287/mnsc.2018.3269

Acquisti, A., & Gross, R. (2006). Imagined communities: Awareness, information sharing, and privacy on the Facebook. Privacy Enhancing Technologies, PET 2006, Berlin, Heidelberg.

Acquisti, A., & Grossklags, J. (2012). An online survey experiment on ambiguity and privacy. *Communications & Strategies, 1*(88), 19-39.

Acquisti, A., Taylor, C., & Wagman, L. (2016). The economics of privacy. *Journal of Economic*

literature, 54(2), 442-492. https://doi.org/10.1257/jel.54.2.442
Adjerid, I., Acquisti, A., & Loewenstein, G. (2018a). Choice architecture, framing, and cascaded privacy choices. *Management Science, 65*(5), 2267-2290. https://doi.org/10.1287/mnsc.2018.3028
Adjerid, I., Peer, E., & Acquisti, A. (2018b). Beyond the privacy paradox: Objective versus relative risk in privacy decision making. *MIS Quarterly, 42*(2), 465-488. https://doi.org/10.25300/MISQ/2018/14316
Ågerfalk, P. J. (2020). Artificial intelligence as digital agency. *European Journal of Information Systems, 29*(1), 1-8. https://doi.org/10.1080/0960085X.2020.1721947
Aiken, L. S., & West, S. G. (1991). *Multiple regression: Testing and interpreting interactions*. SAGE Publications.
Al-Natour, S., Benbasat, I., & Cenfetelli, R. (2021). Designing online virtual advisors to encourage customer self-disclosure: A theoretical model and an empirical test. *Journal of Management Information Systems, 38*(3), 798-827. https://doi.org/10.1080/07421222.2021.1962595
Al-Natour, S., Cavusoglu, H., Benbasat, I., & Aleem, U. (2020). An empirical investigation of the antecedents and consequences of privacy uncertainty in the context of mobile apps. *Information Systems Research, 31*(4), 1037-1063. https://doi.org/10.1287/isre.2020.0931
Alashoor, T., Keil, M., Smith, J., & McConnell, A. R. (2022). Too tired and in too good of a mood to worry about privacy: Explaining the privacy paradox through the lens of effort level in information processing. *Information Systems Research*, Ahead of print. https://doi.org/10.1287/isre.2022.1182
Aloysius, J., Deck, C., & Farmer, A. (2013). Sequential pricing of multiple products: Leveraging revealed preferences of retail customers online and with auto-id technologies. *Information Systems Research, 24*(2), 372-393. https://doi.org/10.1287/isre.1120.0440
Alter, S. (1978). Development patterns for decision support systems. *MIS Quarterly, 2*(3), 33-42. https://doi.org/10.2307/249176
Altman, I. (1975). *The environment and social behavior: Privacy, personal space, territory, and crowding*. Brooks/Cole Publishing Company.
Altman, I. (1977). Privacy regulation: Culturally universal or culturally specific? *Journal of Social Issues, 33*(3), 66-84. https://doi.org/10.1111/j.1540-4560.1977.tb01883.x
Amazon. (2021). *Alexa, Echo devices, and your privacy*. Amazon. https://www.amazon.com/gp/help/customer/display.html?nodeId=GVP69FUJ48X9DK8V

Anderson, C. L., & Agarwal, R. (2011). The digitization of healthcare: Boundary risks, emotion, and consumer willingness to disclose personal health information. *Information Systems Research, 22*(3), 469-490. https://doi.org/10.1287/isre.1100.0335
Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin, 103*(3), 411-423. https://doi.org/10.1037/0033-2909.103.3.411
Anderson, J. C., & Gerbing, D. W. (1991). Predicting the performance of measures in a confirmatory factor analysis with a pretest assessment of their substantive validities. *Journal of Applied Psychology, 76*(5), 732-740. https://doi.org/10.1037/0021-9010.76.5.732
Angst, C. M., & Agarwal, R. (2009). Adoption of electronic health records in the presence of privacy concerns: The elaboration likelihood model and individual persuasion. *MIS Quarterly, 33*(2), 339-370. https://doi.org/10.2307/20650295
Appen. (2021, June 21). *What is data annotation?* Appen. Retrieved September 6, 2021 from https://appen.com/blog/data-annotation/

Apple. (2021, August). *Expanded protections for children*. Apple. Retrieved September 15, 2021 from https://www.apple.com/child-safety/pdf/Expanded_Protections_for_Children_Technology_Summary.pdf

Asatiani, A., Malo, P., Per Rådberg, N., Penttinen, E., Rinta-Kahila, T., & Salovaara, A. (2021). Sociotechnical envelopment of artificial intelligence: An approach to organizational deployment of inscrutable artificial intelligence systems. *Journal of the Association for Information Systems, 22*(2), 8. https://doi.org/10.17705/1jais.00664

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly, 30*(1), 13-28. https://doi.org/10.2307/25148715

Bagozzi, R. P. (1977). Structural equation models in experimental research. *Journal of Marketing Research, 14*(2), 209-226. https://doi.org/10.1177/002224377701400209

Baird, A., & Maruping, L. M. (2021). The next generation of research on is use: A theoretical framework of delegation to and from agentic is artifacts. *MIS Quarterly, 45*(1), 315-341. https://doi.org/10.25300/MISQ/2021/15882

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*, 191-215. https://doi.org/10.1037/0033-295X.84.2.191

Bandura, A. (1986). *Social foundations of thought and action: A social cogntive perspective*. Prentice Hall.

Bansal, G., & Nah, F. F.-H. (2022). Internet privacy concerns revisited: Oversight from surveillance and right to be forgotten as new dimensions. *Information & Management, 59*(3), 103618. https://doi.org/10.1016/j.im.2022.103618

Bansal, G., Zahedi, F. M., & Gefen, D. (2010). The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. *Decision Support Systems, 49*(2), 138-150. https://doi.org/10.1016/j.dss.2010.01.010

Bansal, G., Zahedi, F. M., & Gefen, D. (2015). The role of privacy assurance mechanisms in building trust and the moderating role of privacy concern. *European Journal of Information Systems, 24*(6), 624-644. https://doi.org/10.1057/ejis.2014.41

Bansal, G., Zahedi, F. M., & Gefen, D. (2016). Do context and personality matter? Trust and privacy concerns in disclosing private information online. *Information & Management, 53*(1), 1-21. https://doi.org/10.1016/j.im.2015.08.001

Barlow, J. B., Warkentin, M., Ormond, D., & Dennis, A. (2018). Don't even think about it! The effects of antineutralization, informational, and normative communication on information security compliance. *Journal of the Association for Information Systems, 19*(8), 689-715. https://doi.org/10.17705/1jais.00506

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology, 51*(6), 1173-1182. https://doi.org/10.1037/0022-3514.51.6.1173

Belanger, F., & Crossler, R. E. (2019). Dealing with digital traces: Understanding protective behaviors on mobile devices. *Journal of Strategic Information Systems, 28*(1), 34-49. https://doi.org/10.1016/j.jsis.2018.11.002

Bélanger, F., & Crossler, R. E. (2011). Privacy in the digital age: A review of information privacy research in information systems. *MIS Quarterly, 35*(4), 1017-1041. https://doi.org/10.2307/41409971

Bélanger, F., & James, T. L. (2020). A theory of multilevel information privacy management for the digital era. *Information Systems Research, 31*(2), 510-536. https://doi.org/10.1287/isre.2019.0900

Benbya, H., Pachidi, S., & Jarvenpaa, S. (2021). Artificial intelligence in organizations: Implications for information systems research. *Journal of the Association for Information Systems, 22*(2), 10. https://doi.org/10.17705/1jais.00662

Benlian, A., Klumpe, J., & Hinz, O. (2020). Mitigating the intrusive effects of smart home assistants by using anthropomorphic design features: A multimethod investigation. *Information Systems Journal, 30*(6), 1010-1042. https://doi.org/10.1111/isj.12243

Boh, W. F., & Yellin, D. (2006). Using enterprise architecture standards in managing information

technology. *Journal of Management Information Systems, 23*(3), 163-207. https://doi.org/10.2753/MIS0742-1222230307

Borges, A. F. S., Laurindo, F. J. B., Spínola, M. M., Gonçalves, R. F., & Mattos, C. A. (2021). The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. *International Journal of Information Management, 57*, 102225. https://doi.org/10.1016/j.ijinfomgt.2020.102225

Boss, S. R., Galletta, D. F., Lowry, P. B., Moody, G. D., & Polak, P. (2015). What do users have to fear? Using fear appeals to engender threats and fear that motivate protective security behaviors. *MIS Quarterly, 39*(4), 837-864. https://doi.org/10.25300/MISQ/2015/39.4.5

Boudreau, M.-C., Gefen, D., & Straub, D. W. (2001). Validation in information systems research: A state-of-the-art assessment. *MIS Quarterly, 25*(1), 1-16. https://doi.org/10.2307/3250956

Bourtoule, L., Chandrasekaran, V., Choquette-Choo, C. A., Jia, H., Travers, A., Zhang, B., Lie, D., & Papernot, N. (2021, 24-27 May). Machine unlearning. 2021 IEEE Symposium on Security and Privacy, San Francisco, CA.

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2012). Misplaced confidences: Privacy and the control paradox *Social Psychological and Personality Science, 4*(3), 340-347. https://doi.org/10.1177/1948550612455931

Breward, M., Hassanein, K., & Head, M. (2017). Understanding consumers' attitudes toward controversial information technologies: A contextualization approach. *Information Systems Research, 28*(4), 760-774. https://doi.org/10.1287/isre.2017.0706

Brooks, S., Garcia, M., Lefkovitz, N., Lightman, S., & Nadeau, E. (2017). An introduction to privacy engineering and risk management in Federal systems. *National Institute of Standards and Technology Internal Report 8062*, 1-41. https://doi.org/10.6028/NIST.IR.8062

Brophy, J. (2004). *Motivating students to learn*. Lawrence Erlbaum Associates, Publishers.

Buckman, J. R., Bockstedt, J. C., & Hashim, M. J. (2019). Relative privacy valuations under varying disclosure characteristics. *Information Systems Research, 30*(2), 375-388. https://doi.org/10.1287/isre.2018.0818

Buhrmester, M., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: A new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6*(1), 3-5. https://doi.org/10.1177/1745691610393980

Buxmann, P., Hess, T., & Thatcher, J. B. (2021). AI-based information systems. *Business & Information Systems Engineering, 63*(1), 1-4. https://doi.org/10.1007/s12599-020-00675-8

Cao, Y., & Yang, J. (2015, 17-21 May). Towards making systems forget with machine unlearning. 2015 IEEE Symposium on Security and Privacy, San Jose, CA.

Cavusoglu, H., Phan, T. Q., Cavusoglu, H., & Airoldi, E. M. (2016). Assessing the impact of granular privacy controls on content sharing and disclosure on Facebook. *Information Systems Research, 27*(4), 848-879. https://doi.org/10.1287/isre.2016.0672

Chaudhuri, A., & Holbrook, M. B. (2001). The chain of effects from brand trust and brand affect to brand performance: The role of brand loyalty. *Journal of Marketing, 65*(2), 81-93. https://doi.org/10.1509/jmkg.65.2.81.18255

Cheikh-Ammar, M. (2020). The bittersweet escape to information technology: An investigation of the stress paradox of social network sites. *Information & Management, 57*(8), 103368. https://doi.org/10.1016/j.im.2020.103368

Chen, M., Zhang, Z., Wang, T., Backes, M., Humbert, M., & Zhang, Y. (2021a, November 15–19). When machine unlearning jeopardizes privacy. CCS '21: Proceedings of the 2021 ACM SIGSAC Conference on Computer and Communications Security, Republic of Korea.

Chen, R. (2013). Member use of social networking sites—An empirical examination. *Decision Support Systems, 54*(3), 1219-1227. https://doi.org/10.1016/j.dss.2012.10.028

Chen, R., Kim, D. J., & Rao, H. R. (2021b). A study of social networking site use from a three-pronged security and privacy threat assessment perspective. *Information & Management, 58*(5), 103486. https://doi.org/10.1016/j.im.2021.103486

Chen, Y., Galletta, D. F., Lowry, P. B., Luo, X., Moody, G. D., & Willison, R. (2021c). Understanding inconsistent employee compliance with information security policies through the lens of the extended parallel process model. *Information Systems Research, 32*(3), 1043-1065. https://doi.org/10.1287/isre.2021.1014

Cheng, X., Hou, T., & Mou, J. (2021). Investigating perceived risks and benefits of information privacy disclosure in IT-enabled ride-sharing. *Information & Management, 58*(6), 103450. https://doi.org/10.1016/j.im.2021.103450

Chin, W. W., Thatcher, J. B., Wright, R. T., & Steel, D. (2013). Controlling for common method variance in PLS analysis: The measured latent marker variable approach. New Perspectives in Partial Least Squares and Related Methods, New York, NY.

Choi, B., Wu, Y., Yu, J., & Land, L. (2018). Love at first sight: The interplay between privacy dispositions and privacy calculus in online social connectivity management. *Journal of the Association for Information Systems, 19*(3), 124-151. https://doi.org/10.17705/1jais.00487

Choi, B. C. F., Jiang, Z., Xiao, B., & Kim, S. S. (2015). Embarrassing exposures in online social networks: An integrated perspective of privacy invasion and relationship bonding. *Information Systems Research, 26*(4), 675-694. https://doi.org/10.1287/isre.2015.0602

Choi, B. C. F., Kim, S. S., & Jiang, Z. (2016). Influence of firm's recovery endeavors upon privacy breach on online customer behavior. *Journal of Management Information Systems, 33*(3), 904-933. https://doi.org/10.1080/07421222.2015.1138375

Choi, B. C. F., & Land, L. (2016). The effects of general privacy concerns and transactional privacy concerns on Facebook apps usage. *Information & Management, 53*(7), 868-877. https://doi.org/10.1016/j.im.2016.02.003

Cichy, P., Salge, T. O., & Kohli, R. (2021). Privacy concerns and data sharing in the Internet of Things: Mixed methods evidence from connected cars. *MIS Quarterly, 45*(4), 1863-1891. https://doi.org/10.25300/MISQ/2021/14165

Clarke, R. (1997). *Introduction to dataveillance and information privacy, and definitions of terms*. Xamax Consultancy. http://www.rogerclarke.com/DV/Intro.html

Collins, C., Dennehy, D., Conboy, K., & Mikalef, P. (2021). Artificial intelligence in information systems research: A systematic literature review and research agenda. *International Journal of Information Management, 60*, 102383. https://doi.org/10.1016/j.ijinfomgt.2021.102383

Conger, S., Pratt, J. H., & Loch, K. D. (2013). Personal information privacy and emerging technologies. *Information Systems Journal, 23*(5), 401-417. https://doi.org/10.1111/j.1365-2575.2012.00402.x

Cox, J. (2019, August 7). *Revealed: Microsoft contractors are listening to some Skype calls*. Motherboard. https://www.vice.com/en/article/xweqbq/microsoft-contractors-listen-to-skype-calls

Crossler, R. E., & Bélanger, F. (2019). Why would I use location-protective settings on my smartphone? Motivating protective behaviors and the existence of the privacy knowledge–belief gap. *Information Systems Research, 30*(3), 995-1006. https://doi.org/10.1287/isre.2019.0846

Crossler, R. E., & Posey, C. (2017). Robbing Peter to pay Paul: Surrendering privacy for security's sake in an identity ecosystem. *Journal of the Association for Information Systems, 18*(7), 487-515. https://doi.org/10.17705/1jais.00463

Culnan, M. J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science, 10*(1), 104-115. https://doi.org/10.1287/orsc.10.1.104

Culnan, M. J., & Bies, R. J. (2003). Consumer privacy: Balancing economic and justice considerations.

Journal of Social Issues, 59(2), 323-342. https://doi.org/10.1111/1540-4560.00067
Datta, P., & Chatterjee, S. (2008). The economics and psychology of consumer trust in intermediaries in electronic markets: The EM-trust framework. *European Journal of Information Systems, 17*(1), 12-28. https://doi.org/10.1057/palgrave.ejis.3000729
Davidson, E., Baird, A., & Prince, K. (2018). Opening the envelope of health care information systems research. *Information and Organization, 28*(3), 140-151. https://doi.org/10.1016/j.infoandorg.2018.07.001
Day, M., Turner, G., & Drozdiak, N. (2019, April 10). *Amazon workers are listening to what you tell Alexa*. Bloomberg. https://www.bloomberg.com/news/articles/2019-04-10/is-anyone-listening-to-you-on-alexa-a-global-team-reviews-audio

de Corbiere, F., & Rowe, F. (2013). From ideal data synchronization to hybrid forms of interconnections: Architectures, processes, and data. *Journal of the Association for Information Systems, 14*(10), 550-584. https://doi.org/10.17705/1jais.00345
De Cremer, D., & Kasparov, G. (2021). The ethical AI—paradox: Why better technology needs more and not less human responsibility. *AI and Ethics, 2*, 1-4. https://doi.org/10.1007/s43681-021-00075-y
DeCharms, R. (1968). *Personal causation: The internal affective determinants of behavior*. Academic Press.
Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions *Information Systems Research, 17*(1), 61-80. https://doi.org/10.1287/isre.1060.0080
Dinev, T., Hart, P., & Mullen, M. R. (2008). Internet privacy concerns and beliefs about government surveillance - An empirical investigation. *Journal of Strategic Information Systems, 17*(3), 214-233. https://doi.org/10.1016/j.jsis.2007.09.002
Dinev, T., McConnell, A. R., & Smith, H. J. (2015). Informing privacy research through information systems, psychology, and behavioral economics: Thinking outside the "APCO" box. *Information Systems Research, 26*(4), 639-655. https://doi.org/10.1287/isre.2015.0600
Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts. *European Journal of Information Systems, 22*(3), 295-316. https://doi.org/10.1057/ejis.2012.23
Dowling, G. R., & Staelin, R. (1994). A model of perceived risk and intended risk-handling activity. *Journal of Consumer Research, 21*(1), 119-134. https://doi.org/10.1086/209386
Draper, N. A., & Turow, J. (2019). The corporate cultivation of digital resignation. *New Media & Society, 21*(8), 1824-1839. https://doi.org/10.1177/1461444819833331
Du, R. Y., Netzer, O., Schweidel, D. A., & Mitra, D. (2020). Capturing marketing information to fuel growth. *Journal of Marketing, 85*(1), 163-183. https://doi.org/10.1177/0022242920969198
Facebook. (2021, January 11). *Privacy policy*. Facebook. https://www.facebook.com/privacy/policy/

Fang, X., Hu, P. J.-H., Chau, M., Hu, H.-F., Yang, Z., & Sheng, O. R. L. (2012). A data-driven approach to measure web site navigability. *Journal of Management Information Systems, 29*(2), 173-212. https://doi.org/10.2753/MIS0742-1222290207
Felt, A. P., Ha, E., Egelman, S., Haney, A., Chin, E., & Wagner, D. (2012). *Android permissions: User attention, comprehension, and behavior* Proceedings of the Eighth Symposium on Usable Privacy and Security, Washington, D.C. https://doi.org/10.1145/2335356.2335360

Fernando Libaque-Saenz, C., Wong, S. F., Chang, Y., & Bravo, E. R. (2021). The effect of fair information practices and data collection methods on privacy-related behaviors: A study of mobile apps. *Information & Management, 58*(1), 103284. https://doi.org/10.1016/j.im.2020.103284
Finn, R. L., Wright, D., & Friedewald, M. (2013). Seven types of privacy. In S. Gutwirth, R. Leenes, P. de Hert, & Y. Poullet (Eds.), *European Data Protection: Coming of Age* (pp. 3-32). Springer Netherlands. https://doi.org/10.1007/978-94-007-5170-5_1

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. https://doi.org/10.2307/3151312

Frymier, A. B., Shulman, G. M., & Houser, M. (1996). The development of a learner empowerment measure. *Communication Education, 45*(3), 181-199. https://doi.org/10.1080/03634529609379048

Fügener, A., Grahl, J., Gupta, A., & Ketter, W. (2021). Will humans-in-the-loop become borgs? Merits and pitfalls of working with AI. *MIS Quarterly, 45*(3b), 1527-1556. https://doi.org/10.25300/MISQ/2021/16553

Furneaux, B., & Wade, M. (2017). Impediments to information systems replacement: A calculus of discontinuance. *Journal of Management Information Systems, 34*(3), 902-932. https://doi.org/10.1080/07421222.2017.1373013

Galbreth, M. R., & Shor, M. (2010). The impact of malicious agents on the enterprise software industry. *MIS Quarterly, 34*(3), 595-612. https://doi.org/10.2307/25750693

Gardner, P. L., & Tamir, P. (1989). Interest in biology. Part I: A multidimensional construct. *Journal of Research in Science Teaching, 26*(5), 409-423. https://doi.org/10.1002/tea.3660260506

Ge, R., Zheng, Z., Tian, X., & Liao, L. (2021). Human–robot interaction: When investors adjust the usage of robo-advisors in peer-to-peer lending. *Information Systems Research, 32*(3), 774-785. https://doi.org/10.1287/isre.2021.1009

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Inexperience and experience with online stores: the importance of TAM and trust. *IEEE Transactions on Engineering Management, 50*(3), 307-321. https://doi.org/10.1109/TEM.2003.817277

Gefen, D., & Straub, D. (2005). A practical guide to factorial validity using PLS-Graph: Tutorial and annotated example. *Communications of the Association for Information systems, 16*(1), Article 5. https://doi.org/10.17705/1CAIS.01605

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural equation modeling and regression: Guidelines for research practice. *Communications of the Association for Information systems, 4*(1), 1-79. https://doi.org/10.17705/1CAIS.00407

General Assembly of Virginia. (2021). *Consumer data protection act*. https://lis.virginia.gov/cgi-bin/legp604.exe?212+ful+SB1392ER+pdf

Gerlach, J., Widjaja, T., & Buxmann, P. (2015). Handle with care: How online social network providers' privacy policies impact users' information sharing behavior. *Journal of Strategic Information Systems, 24*(1), 33-43. https://doi.org/10.1016/j.jsis.2014.09.001

Gerlach, J. P., Buxmann, P., & Dinev, T. (2019). "They're all the same!" Stereotypical thinking and systematic errors in users' privacy-related judgments about online services. *Journal of the Association for Information Systems, 20*(6), 787-823. https://doi.org/10.17705/1jais.00551

Google. (2021). *Right to be forgotten overview*. Google. https://support.google.com/legal/answer/10769224?hl=en

Grant, T. D., & Wischik, D. J. (2020). Show us the data: Privacy, explainability, and why the law can't have both. *George Washington Law Review, 88*(6), 1350-1420. https://doi.org/10.17863/CAM.58412

Gregory, R. W., Henfridsson, O., Kaganer, E., & Kyriakou, H. (2020). The role of artificial intelligence and data network effects for creating user value. *Academy of Management Review, 46*(3), 534-551. https://doi.org/10.5465/amr.2019.0178

Griffin, D., & Tversky, A. (1992). The weighing of evidence and the determinants of confidence. *Cognitive Psychology, 24*(3), 411-435. https://doi.org/10.1016/0010-0285(92)90013-R

Grønsund, T., & Aanestad, M. (2020). Augmenting the algorithm: Emerging human-in-the-loop work configurations. *Journal of Strategic Information Systems, 29*(2), 101614. https://doi.org/10.1016/j.jsis.2020.101614

Gu, J., Xu, Y., Xu, H., Zhang, C., & Ling, H. (2017). Privacy concerns for mobile app download: An elaboration likelihood model perspective. *Decision Support Systems, 94*, 19-28. https://doi.org/10.1016/j.dss.2016.10.002

Hair, J. F., Ringle, C. M., & Sarstedt, M. (2011). PLS-SEM: Indeed a silver bullet. *Journal of Marketing Theory and Practice, 19*(2), 139-152. https://doi.org/10.2753/MTP1069-6679190202

Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An assessment of the use of partial least squares structural equation modeling in marketing research. *Journal of the Academy of Marketing Science, 40*(3), 414-433. https://doi.org/10.1007/s11747-011-0261-6

Hair, J. F., Tatham, R. L., Anderson, R. E., & Black, W. (2006). *Multivariate data analysis* (Vol. 6). Pearson Prentice Hall.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*, 316-330. https://doi.org/10.1037/0022-0663.92.2.316

Harman, H. H. (1967). *Modern factor analysis*. University of Chicago Press.

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs, 76*(4), 408-420. https://doi.org/10.1080/03637750903310360

Henseler, J., Hubona, G., & Ray, P. A. (2016). Using PLS path modeling in new technology research: Updated guidelines. *Industrial Management & Data Systems, 116*(1), 2-20. https://doi.org/10.1108/IMDS-09-2015-0382

Herath, T., Chen, R., Wang, J., Banjara, K., Wilbur, J., & Rao, H. R. (2014). Security services as coping mechanisms: An investigation into user intention to adopt an email authentication service. *Information Systems Journal, 24*(1), 61-84. https://doi.org/10.1111/j.1365-2575.2012.00420.x

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111-127. https://doi.org/10.1207/s15326985ep4102_4

Hoehle, H., Aloysius, J. A., Goodarzi, S., & Venkatesh, V. (2019). A nomological network of customers' privacy perceptions: Linking artifact design to shopping efficiency. *European Journal of Information Systems, 28*(1), 91-113. https://doi.org/10.1080/0960085X.2018.1496882

Hoehle, H., & Venkatesh, V. (2015). Mobile application usability: Conceptualization and instrument development. *MIS Quarterly, 39*(2), 435-472. https://doi.org/10.25300/misq/2015/39.2.08

Hoehle, H., Zhang, X., & Venkatesh, V. (2015). An espoused cultural perspective to understand continued intention to use mobile applications: A four-country study of mobile social media application usability. *European Journal of Information Systems, 24*(3), 337-359. https://doi.org/10.1057/ejis.2014.43

Hong, W., Chan, F. K. Y., & Thong, J. Y. L. (2021). Drivers and inhibitors of internet privacy concern: A multidimensional development theory perspective. *Journal of Business Ethics, 168*(3), 539-564. https://doi.org/10.1007/s10551-019-04237-1

Hong, W., & Thong, J. Y. L. (2013). Internet privacy concerns: An integrated conceptualization and four empirical studies. *MIS Quarterly, 37*(1), 275-298. https://doi.org/10.25300/MISQ/2013/37.1.12

Howell, L. C. (2021). Alexa hears with her little ears—But does she have the privilege? *St. Mary's Law Journal, 52*(3), 837-865.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. https://doi.org/10.1080/10705519909540118

Hu, T., Kettinger, W. J., & Poston, R. S. (2015). The effect of online social value on satisfaction and continued use of social media. *European Journal of Information Systems, 24*(4), 391-410. https://doi.org/10.1057/ejis.2014.22

Huang, N., Mojumder, P., Sun, T., Lv, J., & Golden, J. M. (2021). Not registered? Please sign up first: A randomized field experiment on the ex ante registration request. *Information Systems Research, 32*(3), 914-931. https://doi.org/10.1287/isre.2021.0999

Hui, K.-L., Teo, H. H., & Lee, S.-Y. T. (2007). The value of privacy assurance: An exploratory field experiment. *MIS Quarterly, 31*(1), 19-33. https://doi.org/10.2307/25148779

Hulland, J., & Miller, J. (2018). "Keep on turkin'"? *Journal of the Academy of Marketing Science, 46*(5), 789-794. https://doi.org/10.1007/s11747-018-0587-4

James, T., Nottingham, Q., & Kim, B. C. (2013). Determining the antecedents of digital security practices in the general public dimension. *Information Technology and Management, 14*(2), 69-89. https://doi.org/10.1007/s10799-012-0147-4

James, T. L., Wallace, L., Warkentin, M., Kim, B. C., & Collignon, S. E. (2017). Exposing others' information on online social networks (OSNs): Perceived shared risk, its determinants, and its influence on OSN privacy control use. *Information & Management, 54*(7), 851-865. https://doi.org/10.1016/j.im.2017.01.001

James, T. L., Warkentin, M., & Collignon, S. E. (2015). A dual privacy decision model for online social networks. *Information & Management, 52*(8), 893-908. https://doi.org/10.1016/j.im.2015.07.010

James, T. L., Ziegelmayer, J. L., Scott, A. S., & Fox, G. (2021). A multiple-motive heuristic-systematic model for examining how users process Android data and service access notifications. *DATA BASE for Advances in Information Systems, 52*(1), 91–122. https://doi.org/10.1145/3447934.3447941

Jasso, G. (2006). Factorial survey methods for studying beliefs and judgments. *Sociological Methods & Research, 34*(3), 334-423. https://doi.org/10.1177/0049124105283121

Jia, R., Steelman, Z. R., & Jia, H. H. (2022). What makes one intrinsically interested in it? An exploratory study on influences of autistic tendency and gender in the us and India. *MIS Quarterly, 46*(3), 1603-1634. https://doi.org/10.25300/MISQ/2022/16362

Jiang, Z., Heng, C. S., & Choi, B. C. F. (2013). Privacy concerns and privacy-protective behavior in synchronous online social interactions. *Information Systems Research, 24*(3), 579-595. https://doi.org/10.1287/isre.1120.0441

Joshi, A., Kale, S., Chandel, S., & Pal, D. K. (2015). Likert scale: Explored and explained. *British Journal of Applied Science & Technology, 7*(4), 396-403. https://doi.org/10.9734/BJAST/2015/14975

Junglas, I. A., Johnson, N. A., & Spitzmueller, C. (2008). Personality traits and concern for privacy: An empirical study in the context of location-based services. *European Journal of Information Systems, 17*(4), 387-402. https://doi.org/10.1057/ejis.2008.29

Kane, G. C., Young, A. G., Majchrzak, A., & Ransbotham, S. (2021). Avoiding an oppressive future of machine learning: A design theory for emancipatory assistants. *MIS Quarterly, 45*(1), 371-396. https://doi.org/10.25300/MISQ/2021/1578

Karwatzki, S., Dytynko, O., Trenz, M., & Veit, D. (2017a). Beyond the personalization-privacy paradox: Privacy valuation, transparency features, and service personalization. *Journal of Management Information Systems, 34*(2), 369-400. https://doi.org/10.1080/07421222.2017.1334467

Karwatzki, S., Trenz, M., Tuunainen, V. K., & Veit, D. (2017b). Adverse consequences of access to individuals' information: An analysis of perceptions and the scope of organisational influence. *European Journal of Information Systems, 26*(6), 688-715. https://doi.org/10.1057/s41303-017-0064-z

Karwatzki, S., Trenz, M., & Veit, D. (2022). The multidimensional nature of privacy risks: Conceptualisation, measurement and implications for digital services. *Information Systems Journal, 32*(6), 1126-1157. https://doi.org/10.1111/isj.12386

Kaufman, D. J., Murphy-Bollinger, J., Scott, J., & Hudson, K. L. (2009). Public opinion about the importance of privacy in biobank research. *The American Journal of Human Genetics, 85*(5), 643-654. https://doi.org/10.1016/j.ajhg.2009.10.002

Kehr, F., Kowatsch, T., Wentzel, D., & Fleisch, E. (2015). Blissfully ignorant: The effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Information Systems Journal, 25*(6), 607-635. https://doi.org/10.1111/isj.12062

Keith, M. J., Babb, J. S., Lowry, P. B., Furner, C. P., & Abdullat, A. (2015). The role of mobile-computing self-efficacy in consumer information disclosure. *Information Systems Journal, 25*(6), 637-667. https://doi.org/10.1111/isj.12082

Kellogg, K. C., Valentine, M. A., & Christin, A. (2019). Algorithms at work: The new contested terrain of control. *Academy of Management Annals, 14*(1), 366-410. https://doi.org/10.5465/annals.2018.0174

Kennedy, C., Popky, D., & Keeter, S. (2023, April 19). *How public polling has changed in the 21st Century*. Pew Research Center. https://www.pewresearch.org/methods/2023/04/19/how-public-polling-has-changed-in-the-21st-century/

Kim, D. J., Ferrin, D. L., & Rao, H. R. (2008). Trust and satisfaction, two stepping stones for successful e-commerce relationships: A longitudinal exploration. *Information Systems Research, 20*(2), 237-257. https://doi.org/10.1287/isre.1080.0188

Koh, B., Raghunathan, S., & Nault, B. R. (2020). An empirical examination of voluntary profiling: Privacy and quid pro quo. *Decision Support Systems, 132*, 113285. https://doi.org/10.1016/j.dss.2020.113285

Koohikamali, M., Gerhart, N., & Mousavizadeh, M. (2015). Location disclosure on LB-SNAs: The role of incentives on sharing behavior. *Decision Support Systems, 71*, 78-87. https://doi.org/10.1016/j.dss.2015.01.008

Kordzadeh, N., & Ghasemaghaei, M. (2021). Algorithmic bias: Review, synthesis, and future research directions. *European Journal of Information Systems, 31*(3), 388-409. https://doi.org/10.1080/0960085X.2021.1927212

Kordzadeh, N., & Warren, J. (2017). Communicating personal health information in virtual health communities: An integration of privacy calculus model and affective commitment. *Journal of the Association for Information Systems, 18*(1), 45-81. https://doi.org/10.17705/1jais.00446

Krasnova, H., Spiekermann, S., Koroleva, K., & Hildebrand, T. (2010). Online social networks: Why we disclose. *Journal of Information Technology, 25*(2), 109-125. https://doi.org/10.1057/jit.2010.6

Kummer, T.-F., Ryschka, S., & Bick, M. (2018). Why do we share where we are? The influence of situational factors on the conditional value of check-in services. *Decision Support Systems, 115*, 1-12. https://doi.org/10.1016/j.dss.2018.08.012

Kwak, D.-H., Holtkamp, P., & Kim, S. S. (2019). Measuring and controlling social desirability bias: Applications in information systems research. *Journal of the Association for Information Systems, 20*(4), 317-345. https://doi.org/10.17005/1.jais.00537

Lassen, A., Bruselius-Jensen, M., Sommer, H. M., Thorsen, A. V., & Trolle, E. (2007). Factors influencing participation rates and employees' attitudes toward promoting healthy eating at blue-collar worksites. *Health Education Research, 22*(5), 727-736. https://doi.org/10.1093/her/cyl153

Laufer, R. S., & Wolfe, M. (1977). Privacy as a concept and a social issue: A multidimensional developmental theory. *Journal of Social Issues, 33*, 22-42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x

Lebovitz, S., Levina, N., & Lifshitz-Assaf, H. (2021). Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what. *MIS Quarterly, 45*(3b), 1501-1525. https://doi.org/10.25300/MISQ/2021/16564

Lee, D.-J., Ahn, J.-H., & Bang, Y. (2011a). Managing consumer privacy concerns in personalization: A strategic analysis of privacy protection. *MIS Quarterly, 35*(2), 423-444. https://doi.org/10.2307/23044050

Lee, H. C. B., Ba, S., Li, X., & Stallaert, J. (2018). Salience bias in crowdsourcing contests. *Information Systems Research, 29*(2), 401-418. https://doi.org/10.1287/isre.2018.0775

Lee, Y. J., Kauffman, R. J., & Sougstad, R. (2011b). Profit-maximizing firm investments in customer information security. *Decision Support Systems, 51*(4), 904-920. https://doi.org/10.1016/j.dss.2011.02.009

Leidner, D. E., & Tona, O. (2021). The CARE theory of dignity amid personal data digitalization. *MIS Quarterly, 45*(1), 343-370. https://doi.org/10.25300/MISQ/2021/15941

Li, H., Luo, X., Zhang, J., & Xu, H. (2017). Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviors. *Information & Management, 54*(8), 1012-

1022. https://doi.org/10.1016/j.im.2017.02.005
Li, K., Lin, Z., & Wang, X. (2015). An empirical analysis of users' privacy disclosure behaviors on social network sites. *Information & Management, 52*(7), 882-891. https://doi.org/10.1016/j.im.2015.07.006
Li, S. S., & Karahanna, E. (2015). Online recommendation systems in a B2C e-commerce context: A review and future directions. *Journal of the Association for Information Systems, 16*(2), 72-107. https://doi.org/10.17705/1jais.00389
Li, T., & Unger, T. (2012). Willing to pay for quality personalization? Trade-off between quality and privacy. *European Journal of Information Systems, 21*(6), 621-642. https://doi.org/10.1057/ejis.2012.13
Li, Y. (2012). Theories in online information privacy research: A critical review and an integrated framework. *Decision Support Systems, 54*(1), 471-481. https://doi.org/10.1016/j.dss.2012.06.010
Liang, H., & Xue, Y. (2009). Avoidance of information technology threats: A theoretical perspective. *MIS Quarterly, 33*(1), 71-90. https://doi.org/10.2307/20650279
Lin, J., Carter, L., & Liu, D. (2021). Privacy concerns and digital government: Exploring citizen willingness to adopt the COVIDSafe app. *European Journal of Information Systems, 30*(4), 389-402. https://doi.org/10.1080/0960085X.2021.1920857
Lin, S., & Armstrong, D. J. (2019). Beyond information: The role of territory in privacy management behavior on social networking sites. *Journal of the Association for Information Systems, 20*(4), 434-475. https://doi.org/10.17705/1.jais.00540
Lin, X., Featherman, M., & Sarker, S. (2017). Understanding factors affecting users' social networking site continuance: A gender difference perspective. *Information & Management, 54*(3), 383-395. https://doi.org/10.1016/j.im.2016.09.004
Lindebaum, D., Vesa, M., & den Hond, F. (2019). Insights from "The machine stops" to better understand rational assumptions in algorithmic decision making and its implications for organizations. *Academy of Management Review, 45*(1), 247-263. https://doi.org/10.5465/amr.2018.0181
Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology, 86*(1), 114-121. https://doi.org/10.1037/0021-9010.86.1.114
Liu, B., Pavlou, P. A., & Cheng, X. (2022). Achieving a balance between privacy protection and data collection: A field experimental examination of a theory-driven information technology solution. *Information Systems Research, 33*(1), 203-223. https://doi.org/10.1287/isre.2021.1045
Liu, X., Min, Q., Wu, D., & Liu, Z. (2020). How does social network diversity affect users' lurking intention toward social network services? A role perspective. *Information & Management, 57*(7), 103258. https://doi.org/10.1016/j.im.2019.103258
Liu, Z., Min, Q., Zhai, Q., & Smyth, R. (2016). Self-disclosure in Chinese micro-blogging: A social exchange theory perspective. *Information & Management, 53*(1), 53-63. https://doi.org/10.1016/j.im.2015.08.006
Liu, Z., & Wang, X. (2018). How to regulate individuals' privacy boundaries on social network sites: A cross-cultural comparison. *Information & Management, 55*(8), 1005-1023. https://doi.org/10.1016/j.im.2018.05.006
Liu, Z., Wang, X., Min, Q., & Li, W. (2019). The effect of role conflict on self-disclosure in social network sites: An integrated perspective of boundary regulation and dual process model. *Information Systems Journal, 29*(2), 279-316. https://doi.org/10.1111/isj.12195
Lopatovska, I., Rink, K., Knight, I., Raines, K., Cosenza, K., Williams, H., Sorsche, P., Hirsch, D., Li, Q., & Martinez, A. (2018). Talk to me: Exploring user interactions with the Amazon Alexa. *Journal of Librarianship and Information Science, 51*(4), 984-997. https://doi.org/10.1177/0961000618759414
Lowry, P. B., Cao, J., & Everard, A. (2011). Privacy concerns versus desire for interpersonal awareness in driving the use of self-disclosure technologies: The case of instant messaging in two cultures. *Journal of Management Information Systems, 27*(4), 163-200. https://doi.org/10.2753/MIS0742-

1222270406
Lowry, P. B., D'Arcy, J., Hammer, B., & Moody, G. D. (2016). 'Cargo Cult' science in traditional organization and information systems survey research: A case for using nontraditional methods of data collection, including Mechanical Turk and online panels. *Journal of Strategic Information Systems, 25*(3), 232-240. https://doi.org/10.1016/j.jsis.2016.06.002
Lowry, P. B., Moody, G. D., Galletta, D. F., & Vance, A. (2013). The drivers in the use of online whistle-blowing reporting systems. *Journal of Management Information Systems, 30*(1), 153-189. https://doi.org/10.2753/MIS0742-1222300105
Mac, R. (2021, September 3). *Facebook apologizes after A.I. puts 'primates' label on video of black men*. The New York Times. https://www.nytimes.com/2021/09/03/technology/facebook-ai-race-primates.html

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement and validation procedures in MIS and behavioral research: Integrating new and existing techniques. *MIS Quarterly, 35*(2), 293-334. https://doi.org/10.2307/23044045
MacKinnon, D. (2008). *Introduction to statistical mediation analysis*. Erlbaum.
Maddux, J. E., & Rogers, R. W. (1983). Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. *Journal of Experimental Social Psychology, 19*(5), 469-479. https://doi.org/10.1016/0022-1031(83)90023-9
Mai, B., Menon, N. M., & Sarkar, S. (2010). No free lunch: Price premium for privacy seal-bearing vendors. *Journal of Management Information Systems, 27*(2), 189-212. https://doi.org/10.2753/MIS0742-1222270206
Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information Systems Research, 15*(4), 336-355. https://doi.org/10.1287/isre.1040.0032
Malhotra, N. K., Kim, S. S., & Patil, A. (2006). Common method variance in IS research: A comparison of alternative approaches and a reanalysis of past research. *Management Science, 52*(12), 1865-1883. https://doi.org/10.1287/mnsc.1060.0597
Marabelli, M., Newell, S., & Handunge, V. (2021). The lifecycle of algorithmic decision-making systems: Organizational choices and ethical challenges. *Journal of Strategic Information Systems, 30*(3), 101683. https://doi.org/10.1016/j.jsis.2021.101683
Maris, E., Libert, T., & Henrichsen, J. R. (2020). Tracking sex: The implications of widespread sexual data leakage and tracking on porn websites. *New Media & Society, 22*(11), 2018-2038. https://doi.org/10.1177/1461444820924632
Markus, M. L. (2017). Datification, organizational strategy, and IS research: What's the score? *Journal of Strategic Information Systems, 26*(3), 233-241. https://doi.org/10.1016/j.jsis.2017.08.003
Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First- and higher order factor models and their invariance across groups. *Psychological Bulletin, 97*, 562-582. https://doi.org/10.1037/0033-2909.97.3.562
Mason, R. O. (1986). Four ethical issues of the Information Age. *MIS Quarterly, 10*(1), 5-12. https://doi.org/10.2307/248873
Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods, 44*(1), 1-23. https://doi.org/10.3758/s13428-011-0124-6
McDonald, A. M., & Cranor, L. F. (2008). The cost of reading privacy policies. *I/S: A Journal of Law and Policy for the Information Society, 4*(3), 543–568.
McKnight, D. H., Choudhury, V., & Kacmar, C. (2002). Developing and validating trust measures for e-commerce: An integrative typology. *Information Systems Research, 13*(3), 334-359. https://doi.org/10.1287/isre.13.3.334.81
McKnight, D. H., Cummings, L. L., & Chervany, N. L. (1998). Initial trust formation in new organizational relationships. *Academy of Management Review, 23*(3), 473-490. https://doi.org/10.5465/amr.1998.926622

Menard, P., & Bott, G. J. (2020). Analyzing IOT users' mobile device privacy concerns: Extracting privacy permissions using a disclosure experiment. *Computers & Security, 95*, 101856. https://doi.org/10.1016/j.cose.2020.101856

Merrill, M. (2020). An uneasy love triangle between Alexa, your personal life, and data security: Exploring privacy in the digital new age. *Mercer Law Review, 71*(2), 637-658.

Mettler, T., & Wulf, J. (2019). Physiolytics at the workplace: Affordances and constraints of wearables use from an employee's perspective. *Information Systems Journal, 29*(1), 245-273. https://doi.org/10.1111/isj.12205

Metz, C. (2023, May 5). *'The godfather of A.I.' leaves Google and warns of danger ahead*. New York Times. https://www.nytimes.com/2023/05/01/technology/ai-google-chatbot-engineer-quits-hinton.html

Miller, B. K., & Chiodo, B. (2008). Academic entitlement: An adaptation of the Equity Preference Questionnaire for a university setting. Annual Southern Management Association Conference, St. Pete Beach, FL.

Milne, G. R., & Culnan, M. J. (2004). Strategies for reducing online privacy risks: Why consumers read (or don't read) online privacy notices. *Journal of Interactive Marketing, 18*(3), 15-29. https://doi.org/10.1002/dir.20009

Milne, G. R., & Gordon, M. E. (1993). Direct mail privacy-efficiency trade-offs within an implied social contract framework. *Journal of Public Policy & Marketing, 12*(2), 206. https://doi.org/10.1177/074391569101200206

Miltgen, C. L., & Smith, H. J. (2015). Exploring information privacy regulation, risks, trust, and behavior. *Information & Management, 52*(6), 741-759. https://doi.org/10.1016/j.im.2015.06.006

Miltgen, C. L., & Smith, H. J. (2019). Falsifying and withholding: Exploring individuals' contextual privacy-related decision-making. *Information & Management, 56*(5), 696-717. https://doi.org/10.1016/j.im.2018.11.004

Mirzaei, T., & Esmaeilzadeh, P. (2021). Engagement in online health communities: Channel expansion and social exchanges. *Information & Management, 58*(1), 103404. https://doi.org/10.1016/j.im.2020.103404

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology, 85*, 424-436. https://doi.org/10.1037/0022-0663.85.3.424

Moody, G. D., Galletta, D. F., & Lowry, P. B. (2014). When trust and distrust collide online: The engenderment and role of consumer ambivalence in online consumer behavior. *Electronic Commerce Research and Applications, 13*(4), 266-282. https://doi.org/10.1016/j.elerap.2014.05.001

Moody, G. D., Lowry, P. B., & Galletta, D. F. (2017). It's complicated: Explaining the relationship between trust, distrust, and ambivalence in online transaction relationships using polynomial regression analysis and response surface analysis. *European Journal of Information Systems, 26*(4), 379-413. https://doi.org/10.1057/s41303-016-0027-9

Moody, G. D., Siponen, M., & Pahnila, S. (2018). Toward a unified model of information security policy compliance. *MIS Quarterly, 42*(1), 285–312. https://doi.org/10.25300/misq/2018/13853

Morgeson, F. P., & Hofmann, D. A. (1999). The structure and function of collective constructs: Implications for multilevel research and theory development. *Academy of Management Review, 24*(2), 249-265. https://doi.org/10.5465/amr.1999.1893935

Moriuchi, E. (2019). Okay, Google!: An empirical study on voice assistants on consumer engagement and loyalty. *Psychology & Marketing, 36*(5), 489-501. https://doi.org/10.1002/mar.21192

Mulligan, D. K., Regan, P. M., & King, J. (2020). The fertile dark matter of privacy takes on the dark patterns of surveillance. *Journal of Consumer Psychology, 30*(4), 767-773. https://doi.org/10.1002/jcpy.1190

Murray, A., Rhymer, J., & Sirmon, D. G. (2020). Humans and technology: Forms of conjoined agency in organizations. *Academy of Management Review, 46*(3), 552-571. https://doi.org/10.5465/amr.2019.0186

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues, 56*(1), 81-103. https://doi.org/10.1111/0022-4537.00153

Nass, C., Moon, Y., Fogg, B. J., & Reeves, B. (1995). Can computer personalities be human personalities? *International Journal of Human-Computer Studies, 43*(2), 223-239. https://doi.org/10.1006/ijhc.1995.1042

Neumann, N., Tucker, C. E., & Whitfield, T. (2019). Frontiers: How effective is third-party consumer profiling? Evidence from field studies. *Marketing Science, 38*(6), 918-926. https://doi.org/10.1287/mksc.2019.1188

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs, 41*(1), 100-126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Obar, J. A., & Oeldorf-Hirsch, A. (2020). The biggest lie on the Internet: Ignoring the privacy policies and terms of service policies of social networking services. *Information, Communication & Society, 23*(1), 128-147. https://doi.org/10.1080/1369118X.2018.1486870

Office of the Attorney General of California. (2021). *California Consumer Privacy Act (CCPA)*. https://oag.ca.gov/privacy/ccpa

Ogbanufe, O., & Gerhart, N. (2020). The mediating influence of smartwatch identity on deep use and innovative individual performance. *Information Systems Journal, 30*(6), 977-1009. https://doi.org/10.1111/isj.12288

Olwal, A., & Dementyev, A. (2022). *Hidden interfaces for ambient computing: Enabling interaction in everyday materials through high-brightness visuals on low-cost matrix displays* Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, New Orleans, LA, USA. https://doi.org/10.1145/3491102.3517674

Ozdemir, Z. D., Smith, H. J., & Benamati, J. H. (2017). Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study. *European Journal of Information Systems, 26*(6), 642-660. https://doi.org/10.1057/s41303-017-0056-z

Parks, R., Xu, H., Chu, C.-H., & Lowry, P. B. (2017). Examining the intended and unintended consequences of organisational privacy safeguards. *European Journal of Information Systems, 26*(1), 37-65. https://doi.org/10.1057/s41303-016-0001-6

Parliament and Council of the European Union. (2016, April 27). *EU General Data Protection Regulation (GDPR): Regulation (EU) 2016/679*. Official Journal of the European Union. https://gdpr-info.eu/

Pavlou, P. A. (2011). State of the information privacy literature: Where are we now and where should we go? *MIS Quarterly, 35*(4), 977-988. https://doi.org/10.2307/41409969

Pavlou, P. A., & Gefen, D. (2004). Building effective online marketplaces with institution-based trust. *Information Systems Research, 15*(1), 37-59. https://doi.org/10.1287/isre.1040.0015

Pelley, S. (2023, April 16). *Is artificial intelligence advancing too quickly? What AI leaders at Google say*. CBS News. https://www.cbsnews.com/news/google-artificial-intelligence-future-60-minutes-transcript-2023-04-16/

Pew Research Center. (2019, November 15). *Americans and privacy: Concerned, confused and feeling lack of control over their personal information*. Pew Research Center,. https://www.pewresearch.org/internet/2019/11/15/americans-and-privacy-concerned-confused-and-feeling-lack-of-control-over-their-personal-information/

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903. https://doi.org/10.1037/0021-9010.88.5.879

Porteous, J. D. (1976). Home: The territorial core. *Geographical Review, 66*(4), 383-390. https://doi.org/10.2307/213649

Preibusch, S. (2013). Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human-Computer Studies, 71*(12), 1133-1143. https://doi.org/10.1016/j.ijhcs.2013.09.002

Raddatz, N., Coyne, J., Menard, P., & Crossler, R. E. Becoming a blockchain user: Understanding consumers' benefits realisation to use blockchain-based applications. *European Journal of Information Systems, 32*(2), 287-314. https://doi.org/10.1080/0960085X.2021.1944823

Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science, 48*(1), 137-141. https://doi.org/10.1007/s11747-019-00710-5

Rai, A., Constantinides, P., & Sarker, S. (2019). Next-generation digital platforms: Toward human–AI hybrids. *MIS Quarterly, 43*(1), iii–x.

Raisch, S., & Krakowski, S. (2021). Artificial intelligence and management: The automation–augmentation paradox. *Academy of Management Review, 46*(1), 192-210. https://doi.org/10.5465/amr.2018.0072

Renninger, K. A., & Hidi, S. E. (2019). Interest development and learning. In K. A. Renninger & S. E. Hidi (Eds.), *The Cambridge Handbook of Motivation and Learning* (pp. 265-290). Cambridge University Press. https://doi.org/10.1017/9781316823279.013

Research and Markets. (2022, Aug. 10). *Digital advertising global market to surpass $1.79 trillion by 2031*. Cision PR Newswire. https://www.prnewswire.com/news-releases/digital-advertising-global-market-to-surpass-1-79-trillion-by-2031--301603429.html

Rheu, M., Shin, J. Y., Peng, W., & Huh-Yoo, J. (2021). Systematic review: Trust-building factors and implications for conversational agent design. *International Journal of Human–Computer Interaction, 37*(1), 81-96. https://doi.org/10.1080/10447318.2020.1807710

Rigdon, E. E. (1996). CFI versus RMSEA: A comparison of two fit indexes for structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 3*(4), 369-379. https://doi.org/10.1080/10705519609540052

Robinson, M. A. (2018). Using multi-item psychometric scales for research and practice in human resource management. *Human Resource Management, 57*(3), 739-750. https://doi.org/10.1002/hrm.21852

Rosenstock, I. M. (1974). Historical origins of the health belief model. *Health Education Monographs, 2*(4), 328-335. https://doi.org/10.1177/109019817400200403

Rungtusanatham, M., Wallin, C., & Eckerd, S. (2011). The vignette in a scenario-based role-playing experiment. *Journal of Supply Chain Management, 47*(3), 9-16. https://doi.org/10.1111/j.1745-493X.2011.03232.x

Sasidharan, S., Santhanam, R., Brass, D. J., & Sambamurthy, V. (2011). The effects of social network structure on enterprise systems success: A longitudinal multilevel analysis. *Information Systems Research, 23*(3-part-1), 658-678. https://doi.org/10.1287/isre.1110.0388

Schneider, B. A., Avivi-Reich, M., & Mozuraitis, M. (2015). A cautionary note on the use of the Analysis of Covariance (ANCOVA) in classification designs with and without within-subject factors. *Frontiers in Psychology, 6*, Article 474. https://doi.org/10.3389/fpsyg.2015.00474

Schraw, G., Bruning, R., & Svoboda, C. (1995). Sources of situational interest. *Journal of Reading Behavior, 27*(1), 1-17. https://doi.org/10.1080/10862969509547866

Schuetz, S., & Venkatesh, V. (2020). The rise of human machines: How cognitive computing systems challenge assumptions of user-system interaction. *Journal of the Association for Information Systems, 21*(2), 460-482. https://doi.org/10.17705/1jais.00608

Schuetz, S. W., Lowry, P. B., Pienta, D. A., & Thatcher, J. B. (2021). Improving the design of

information security messages by leveraging the effects of temporal distance and argument nature. *Journal of the Association for Information Systems, 22*(5), 1376-1428. https://doi.org/10.17705/1jais.00697

Schultz, S., & Shulman, G. (1993). The development and assessment of the job empowerment instrument. Joint Central States Communication Association and Southern States Communication Association annual convention, Lexington, KY.

Schwaig, K. S., Segars, A. H., Grover, V., & Fiedler, K. D. (2013). A model of consumers' perceptions of the invasion of information privacy. *Information & Management, 50*(1), 1-12. https://doi.org/10.1016/j.im.2012.11.002

Segaar, D., Willemsen, M. C., Bolman, C., & De Vries, H. (2007). Nurse adherence to a minimal-contact smoking cessation intervention on cardiac wards. *Research in Nursing & Health, 30*(4), 429-444. https://doi.org/10.1002/nur.20204

Seitz, P. (2021, August 2). *Survey reveals which tech companies consumers trust the most*. Investor's Business Daily. https://www.investors.com/news/technology/tech-stocks-survey-reveals-which-tech-companies-consumers-trust-the-most/

Shen, X.-L., Li, Y.-J., Sun, Y., Chen, Z., & Wang, F. (2019). Understanding the role of technology attractiveness in promoting social commerce engagement: Moderating effect of personal interest. *Information & Management, 56*(2), 294-305. https://doi.org/10.1016/j.im.2018.09.006

Sheng, H., Nah, F. F.-H., & Siau, K. (2008). An experimental study on ubiquitous commerce adoption: Impact of personalization and privacy concerns. *Journal of the Association for Information Systems, 9*(6), 344-377. https://doi.org/10.17705/1jais.00161

Shih, H.-p., Lai, K.-H., & Cheng, T. C. E. (2017). Constraint-based and dedication-based mechanisms for encouraging online self-disclosure: Is personalization the only thing that matters? *European Journal of Information Systems, 26*(4), 432-450. https://doi.org/10.1057/s41303-016-0031-0

Siponen, M., & Vance, A. (2010). Neutralization: New insights into the problem of employee information systems security policy violations. *MIS Quarterly, 34*(3), 487-502. https://doi.org/10.2307/25750688

Smith, H. J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly, 35*(4), 989-1016. https://doi.org/10.2307/41409970

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly, 20*(2), 167-196. https://doi.org/10.2307/249477

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. *Sociological Methodology, 13*, 290-312. https://doi.org/10.2307/270723

Solove, D. J. (2015). The meaning and value of privacy. In B. Roessler & D. Mokrosinska (Eds.), *Social dimensions of privacy: Interdisciplinary perspectives* (pp. 71-82). Cambridge University Press. https://doi.org/10.1017/CBO9781107280557

Solove, D. J. (2021). The myth of the privacy paradox. *George Washington Law Review, 89*(1), 1-51.

Son, J.-Y., & Kim, S. S. (2008). Internet users' information privacy-protective responses: A taxonomy and a nomological model. *MIS Quarterly, 32*(3), 503-529. https://doi.org/10.2307/25148854

Spiekermann, S., Böhme, R., Acquisti, A., & Hui, K.-L. (2015). Personal data markets. *Electronic Markets, 25*(2), 91-93. https://doi.org/10.1007/s12525-015-0190-1

Spiekermann, S., & Korunovska, J. (2017). Towards a value theory for personal data. *Journal of Information Technology, 32*(1), 62-84. https://doi.org/10.1057/jit.2016.4

Spreitzer, G. M. (1995). Psychological empowerment in the workplace: Dimensions, measurement, and validation. *Academy of Management Journal, 38*(5), 1442-1465. https://doi.org/10.5465/256865

Stecklow, S., Cunningham, W., & Jin, H. (2023, April 6). *Special report: Tesla workers shared sensitive images recorded by customer cars*. Reuters. https://www.reuters.com/technology/tesla-workers-shared-sensitive-images-recorded-by-customer-cars-2023-04-06/

Steelman, Z. R., Hammer, B. I., & Limayem, M. (2014). Data collection in the digital age: Innovative alternatives to student samples. *MIS Quarterly, 38*(2), 355-378. https://doi.org/10.25300/MISQ/2014/38.2.02

Stewart, K. A., & Segars, A. H. (2002). An empirical examination of the concern for information privacy instrument. *Information Systems Research, 13*(1), 36-49. https://doi.org/10.1287/isre.13.1.36.97

Stone, E. F., & Stone, D. L. (1990). Privacy in organizations: Theoretical issues, research findings, and protection mechanisms. In K. M. Rowland & G. R. Ferris (Eds.), *Research in personnel and human resources management* (Vol. 8, pp. 349-411). JAI Press.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information systems, 13*(1), 380--427. https://doi.org/10.17705/1cais.01324

Sun, C., Shi, Z., Liu, X., Ghose, A., Li, X., & Xiong, F. (2021a). The effect of voice AI on consumer purchase and search behavior. *Marketing Science Institute Working Paper Series*, 1–43. https://doi.org/10.2139/ssrn.3480877

Sun, Y., Wang, N., & Shen, X.-L. (2021b). Calculus interdependency, personality contingency, and causal asymmetry: Toward a configurational privacy calculus model of information disclosure. *Information & Management, 58*(8), 103556. https://doi.org/10.1016/j.im.2021.103556

Sutanto, J., Palme, E., Tan, C.-H., & Phang, C. W. (2013). Addressing the personalization-privacy paradox: An empirical assessment from a field experiment on smartphone users. *MIS Quarterly, 37*(4), 1141-1164. https://doi.org/10.25300/MISQ/2013/37.4.07

Tang, Y., & Ning, X. (2023). Understanding user misrepresentation behavior on social apps: The perspective of privacy calculus theory. *Decision Support Systems, 165*, 113881. https://doi.org/10.1016/j.dss.2022.113881

Teodorescu, M. H. M., Morse, L., Awwad, Y., & Kane, G. C. (2021). Failures of fairness in automation require a deeper understanding of human–ML augmentation. *MIS Quarterly, 45*(3b), 1483-1499. https://doi.org/10.25300/MISQ/2021/16535

Teubner, T., & Flath, C. M. (2019). Privacy in the sharing economy. *Journal of the Association for Information Systems, 20*(3), 213-242. https://doi.org/10.17705/1jais.00534

Thomas, K. W., & Velthouse, B. A. (1990). Cognitive elements of empowerment: An "interpretive" model of intrinsic task motivation. *Academy of Management Review, 15*(4), 666-681. https://doi.org/10.5465/amr.1990.4310926

Trenz, M., Huntgeburth, J., & Veit, D. (2018). Uncertainty in cloud service relationships: Uncovering the differential effect of three social influence processes on potential and current users. *Information & Management, 55*(8), 971-983. https://doi.org/10.1016/j.im.2018.05.002

Trevino, L. K. (1992). Experimental approaches to studying ethical-unethical behavior in organizations. *Business Ethics Quarterly, 2*(2), 121-136. https://doi.org/10.2307/3857567

Tsai, J. Y., Egelman, S., Cranor, L., & Acquisti, A. (2011). The effect of online privacy information on purchasing behavior: An experimental study. *Information Systems Research, 22*(2), 254-268. https://doi.org/10.1287/isre.1090.0260

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1-10. https://doi.org/10.1007/BF02291170

Turban, E., & Watkins, P. R. (1986). Integrating expert systems and decision support systems. *MIS Quarterly, 10*(2), 121-136. https://doi.org/10.2307/249031

Turel, O., & Qahri-Saremi, H. (2023). Responses to ambivalence toward social networking sites: A typological perspective. *Information Systems Journal, 33*(2), 385-416. https://doi.org/10.1111/isj.12407

Vance, A., Anderson, B. B., Kirwan, C. B., & Eargle, D. (2014). Using measures of risk perception to predict information security behavior: insights from electroencephalography (EEG). *Journal of the Association for Information Systems, 15*(10), 679-722. https://doi.org/10.17705/1jais.00375

Vance, A., Lowry, P. B., & Eggett, D. (2015). Increasing accountability through user-interface design artifacts: A new approach to addressing the problem of access-policy violations. *MIS Quarterly,*

39(2), 345-366. https://doi.org/10.25300/MISQ/2015/39.2.04
Villaronga, E. F., Kieseberg, P., & Li, T. (2018). Humans forget, machines remember: Artificial intelligence and the right to be forgotten. *Computer Law & Security Review*, *34*(2), 304-313. https://doi.org/10.1016/j.clsr.2017.08.007
Vincent, J. (2017, September 13). *The iPhone X's new neural engine exemplifies Apple's approach to AI*. The Verge. https://www.theverge.com/2017/9/13/16300464/apple-iphone-x-ai-neural-engine

Vincent, J. (2021, June 7). *Apple's Siri will finally work without an internet connection with on-device speech recognition*. The Verge. https://www.theverge.com/2021/6/7/22522993/apple-siri-on-device-speech-recognition-no-internet-wwdc

Wakefield, R. (2013). The influence of user affect in online information disclosure. *Journal of Strategic Information Systems*, *22*(2), 157-174. https://doi.org/10.1016/j.jsis.2013.01.003
Wall, J. D., Lowry, P. B., & Barlow, J. B. (2016). Organizational violations of externally governed privacy and security rules: Explaining and predicting selective violations under conditions of strain and excess. *Journal of the Association for Information Systems*, *17*(1), 39-76. https://doi.org/10.17705/1jais.00420
Wang, S.-C., & Wu, J.-H. (2014). Proactive privacy practices in transition: Toward ubiquitous services. *Information & Management*, *51*(1), 93-103. https://doi.org/10.1016/j.im.2013.09.005
Wang, W.-T., Wang, Y.-S., & Liu, E.-R. (2016). The stickiness intention of group-buying websites: The integration of the commitment-trust theory and e-commerce success model. *Information & Management*, *53*(5), 625-642. https://doi.org/10.1016/j.im.2016.01.006
Warkentin, M., Goel, S., & Menard, P. (2017). Shared benefits and information privacy: What determines smart meter technology adoption? *Journal of the Association for Information Systems*, *18*(11), 758-786. https://doi.org/10.17705/1jais.00474
Warren, S. D., & Brandeis, L. D. (1890, December 15). *Right to privacy*. Harvard Law Review.
Wattal, S., Telang, R., Mukhopadhyay, T., & Boatwright, P. (2012). What's in a "Name"? Impact of use of customer information in e-mail advertisements. *Information Systems Research*, *23*(3), 679-697. https://doi.org/10.1287/isre.1110.0384
Weber, K., Martin, M. M., & Cayanus, J. L. (2005). Student interest: A two-study re-examination of the concept. *Communication Quarterly*, *53*(1), 71-86. https://doi.org/10.1080/01463370500055996
Weber, K., & Patterson, B. R. (2000). Student interest, empowerment and motivation. *Communication Research Reports*, *17*(1), 22-29. https://doi.org/10.1080/08824090009388747
Weber, R. (2021). Constructs and indicators: An ontological analysis. *MIS Quarterly*, *45*(4), 1644-1678. https://doi.org/10.25300/MISQ/2021/15999
Wells, E. L., & Marwell, G. (1976). *Self-esteem: Its conceptualization and measurement*. Sage.
Westin, A. F. (1967). *Privacy and freedom*. Athenum.
Westin, A. F. (2000). Intrusions. *Public Perspective*, *11*(6), 8-11.
Whetten, D. A., Felin, T., & King, B. G. (2009). The practice of theory borrowing in organizational studies: Current issues and future directions. *Journal of Management*, *35*(3), 537-563. https://doi.org/10.1177/0149206308330556
Wissner-Gross, A. (2016). *Datasets over algorithms*. Edge. https://www.edge.org/response-detail/26587

Witte, K. (1992). Putting the fear back into fear appeals: The extended parallel process model. *Communication Monographs*, *59*(4), 329-349. https://doi.org/10.1080/03637759209376276
Wottrich, V. M., van Reijmersdal, E. A., & Smit, E. G. (2018). The privacy trade-off for mobile app downloads: The roles of app value, intrusiveness, and privacy concerns. *Decision Support Systems*, *106*, 44-52. https://doi.org/10.1016/j.dss.2017.12.003
Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., & Marett, K. (2014). Influence techniques in phishing attacks: An examination of vulnerability and resistance. *Information Systems Research*, *25*(2), 385-400. https://doi.org/10.1287/isre.2014.0522

Wu, K.-W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior, 28*(3), 889-897. https://doi.org/10.1016/j.chb.2011.12.008

Wu, Z., & Luo, J. (2022). Online information privacy and price: A theoretical model and empirical tests. *Information & Management, 59*(2), 103583. https://doi.org/10.1016/j.im.2021.103583

Xu, C., Peak, D., & Prybutok, V. (2015). A customer value, satisfaction, and loyalty perspective of mobile application recommendations. *Decision Support Systems, 79*, 171-183. https://doi.org/10.1016/j.dss.2015.08.008

Xu, H., Dinev, T., Smith, H. J., & Hart, P. (2008, December 14-17). Examining the formation of individual's privacy concerns: Toward an integrative view. ICIS 2008 Proceedings, Paris, France.

Xu, H., Dinev, T., Smith, J., & Hart, P. (2011a). Information privacy concerns: Linking individual perceptions with institutional privacy assurances. *Journal of the Association for Information Systems, 12*(12), 798-824. https://doi.org/10.17705/1jais.00281

Xu, H., Gupta, S., Rosson, M. B., & Carroll, J. M. (2012a). *Measuring mobile users' concerns for information privacy* Thirty Third International Conference on Information Systems, ICIS 2012, Orlando, FL. https://aisel.aisnet.org/icis2012/proceedings/ISSecurity/10

Xu, H., Luo, X., Carroll, J. M., & Rosson, M. B. (2011b). The personalization privacy paradox: An exploratory study of decision making process for location-aware marketing. *Decision Support Systems, 51*(1), 42-52. https://doi.org/10.1016/j.dss.2010.11.017

Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2009). The role of push-pull technology in privacy calculus: The case of location-based services. *Journal of Management Information Systems, 26*(3), 135-173. https://doi.org/10.2753/MIS0742-1222260305

Xu, H., Teo, H.-H., Tan, B. C. Y., & Agarwal, R. (2012b). Effects of individual self-protection, industry self-regulation, and government regulation on privacy concerns: A study of location-based services. *Information Systems Research, 23*(4), 1342-1363. https://doi.org/10.1287/isre.1120.0416

Yaraghi, N., Gopal, R. D., & Ramesh, R. (2019). Doctors' orders or patients' preferences? Examining the role of physicians in patients' privacy decisions on health information exchange platforms. *Journal of the Association for Information Systems, 20*(7), 928-952. https://doi.org/10.17705/1jais.00557

Yun, H., Lee, G., & Kim, D. J. (2019). A chronological review of empirical research on personal information privacy concerns: An analysis of contexts and research constructs. *Information & Management, 56*(4), 570-601. https://doi.org/10.1016/j.im.2018.10.001

Zeithaml, V. A. (1988). Consumer perceptions of price, quality, and value: A means-end model and synthesis of evidence. *Journal of Marketing, 52*(3), 2-22. https://doi.org/10.2307/1251446

Zhang, N. A., Wang, C. A., Karahanna, E., & Xu, Y. (2022). Peer privacy concern: Conceptualization and measurement. *MIS Quarterly, 46*(1), 491-530. https://doi.org/10.25300/MISQ/2022/14861

Zhang, X., Liu, S., Chen, X., Wang, L., Gao, B., & Zhu, Q. (2018). Health information privacy concerns, antecedents, and information disclosure intention in online health communities. *Information & Management, 55*(4), 482-493. https://doi.org/10.1016/j.im.2017.11.003

Zhu, H., Ou, C. X. J., van den Heuvel, W. J. A. M., & Liu, H. (2017). Privacy calculus and its utility for personalization services in e-commerce: An analysis of consumer decision-making. *Information & Management, 54*(4), 427-437. https://doi.org/10.1016/j.im.2016.10.001

Zhu, Y.-Q., Kanjanamekanant, K., & Chiu, Y.-T. (2023). Reconciling the personalization-privacy paradox: Exploring privacy boundaries in online personalized advertising. *Journal of the Association for Information Systems, 24*(1), 294-316. https://doi.org/10.17705/1jais.00775

Zimmer, J. C., Arsal, R., Al-Marzouq, M., Moore, D., & Grover, V. (2010a). Knowing your customers: Using a reciprocal relationship to enhance voluntary information disclosure. *Decision Support Systems, 48*(2), 395-406. https://doi.org/10.1016/j.dss.2009.10.003

Zimmer, J. C., Arsal, R. E., Al-Marzouq, M., & Grover, V. (2010b). Investigating online information disclosure: Effects of information relevance, trust and risk. *Information & Management, 47*(2), 115-123. https://doi.org/10.1016/j.im.2009.12.003

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology, 30*(1), 75-89. https://doi.org/10.1057/jit.2015.5

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.

---

[i] Although the relationship between HI and C is significant, after controlling for the effects of covariates on willingness to share, the relationship between C and WTS is not significant. Therefore, the mediation effect of HI → C → WTS cannot be inferred.

[ii] We chose ANOVA instead of ANCOVA because various assumptions must be met to perform an ANCOVA and interpret its results accurately (Schneider et al., 2015). First, ANCOVA assumes that a linear relationship is present between the dependent variable and the covariate. Furthermore, the slope of the line relating the dependent variable to the covariate cannot differ across conditions in an experiment. Second, for ANCOVA to be valid for experimentally-defined between-subject factors, then the expected value of the covariate must be the same for all participants across every condition in the experiment. This assumption does not hold when a between-subjects condition is based on a subset of participants who are also classified by other factors (i.e., more than one factor is being manipulated in a single condition). That is, "caution should be employed when considering an ANCOVA when one or more of the between subjects factors are based on a classification of participants into different groups" (Schneider et al., 2015, p. 2). To test treatment effects of a single factor with ANCOVA, the between-subjects conditions for that factor should not include participants who also received different sets of treatments. When this occurs, the expected values of the covariates may not be the same across groups. An ANCOVA should be used to provide a "valid test of the null hypothesis that the relationship of the covariate to the dependent variable is zero" (Schneider et al., 2015, p. 7). In our experiment, this is not the case. Participants are classified into one of 17 conditions and received different sets of treatments. Therefore, we used an ANOVA to detect whether the means of two groups were significantly different for each treatment we administered.