December 2024

Privacy-Utility Trade-Offs in Genetic Data Sharing and the Moderating Role of Social Distance: An Interdependent Privacy Calculus

Scott Thiebes
Karlsruhe Institute of Technology, scott.thiebes@kit.edu

Manuel Schmidt-Kraepelin
Karlsruhe Institute of Technology, manuel.schmidt-kraepelin@kit.edu

Philipp Toussaint
Karlsruhe Institute of Technology, philipp.toussaint@kit.edu

Kalle Lyytinen
Case Western Reserve University, kalle@case.edu

Ali Sunyaev
Karlsruhe Institute of Technology, sunyaev@kit.edu

Privacy-Utility Trade-Offs in Genetic Data Sharing and the Moderating Role of Social Distance: An Interdependent Privacy Calculus

Completed Research Paper

Scott Thiebes
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
scott.thiebes@kit.edu

Manuel Schmidt-Kraepelin
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
manuel.schmidt-kraepelin@kit.edu

Philipp A. Toussaint
Karlsruhe Institute of Technology
76131 Karlsruhe, Germany
philipp.toussaint@kit.edu

Kalle Lyytinen
Case Western Reserve University
Cleveland, Ohio, United States
kalle@case.edu

Ali Sunyaev
Technical University of Munich
74076 Heilbronn, Germany
ali.sunyaev@tum.de

Abstract

The sharing of sensitive data increasingly poses privacy risks not only to the disclosing individuals, but also to their families and friends. Genetic data sharing is a prime example of the tensions between personal and interdependent privacy risks and benefits that individuals must trade off when making disclosure decisions. Surprisingly, extant research has mostly neglected interdependent privacy-utility trade-offs and focused on personal privacy-utility trade-offs in genetic data sharing. We draw on the privacy calculus and the concept of interdependent privacy, as well as on construal level theory to theorize about the effects of personal and interdependent privacy-utility trade-offs in genetic data sharing. To test our hypotheses, we conducted a scenario-based factorial survey using contextualized vignettes in a typical genetic data sharing scenario. While our results support the hypothesized direct effects of personal and interdependent privacy-utility trade-offs, the role of social distance as a moderator within the interdependent calculus remains inconclusive.

Keywords: interdependent privacy, genetic data sharing, privacy calculus, social distance, construal level theory, vignette study

Introduction

In our ever-digitized world, individuals increasingly share data with third parties that, when shared, might not only pose privacy risks to themselves but also privacy risks to their relatives, friends, or acquaintances (Franz and Benlian 2022). A prime example of such interdependent privacy risks (i.e., an individual A's

expectation of losses for an individual B, arising due to individual A disclosing their own personal data to a third party), is the sharing of genetic data.

Genetic data are an increasingly affordable (Thiebes et al. 2020) and very rich source of information that individuals can exploit to learn more about themselves (Roberts et al. 2017). Beneficial information that individuals might learn from their genetic data include disease risks, personal traits, or information about ancestry and family relationships. At the same time, sharing genetic data also poses various threats to an individual's privacy (Anderson and Agarwal 2011). Potential negative consequences of genetic data disclosure include, for example, social marginalization and stigmatization (Green and Botkin 2003), or unsolicited information about genetic predispositions (Middleton et al. 2019). Next to privacy risks for the disclosing individual, sharing personal genetic data may also have serious implications for related others, especially for blood relatives who share parts of their DNA with the disclosing individual. For example, when Joseph James DeAngelo — the infamous Golden State Killer — was charged with first degree murder in 2018, he was the first criminal, in a following series of reopened cases, to be identified decades after his crimes (Erlich et al. 2018). One of DeAngelo's distant relatives had used a direct-to-consumer (DTC) genetic testing service and uploaded their genetic data to a public online database. Authorities matched that data to DNA found at the crime scenes and used it to construct a family tree that eventually led to the identification of DeAngelo as the perpetrator more than 40 years after his first murder. While the sharing of genetic data has benefited society in the case of DeAngelo, avenues for infringing the privacy of blood relatives like those of the original data-sharing individual of genetic data are possible. In a more recent case, hackers breached the data of millions of customers of the popular DTC genetic testing service 23andMe, supposedly targeting especially Jewish and Chinese customers (Carballo et al. 2024). As a result, customers with Jewish heritage, who had originally used the service to connect with relatives, expressed serious concerns that their families could become targets of antisemitic hate speech and violence due to the current conflict between Israel and Gaza (Carballo et al. 2024). In some cases, however, sharing one's own genetic data might also be beneficial to related others. For example, although men have a much lower risk of suffering from breast cancer caused by mutations of the BRCA1/2 genes than women, it is often reasonable for them to undergo genetic screening to complete the family anamnesis for their female descendants and thereby provide valuable information about their health (Smith et al. 2007).

Previous research has extensively investigated individuals' data sharing intentions and decisions in light of apparent privacy-utility trade-offs (i.e., cognitive trade-offs made between the benefits and privacy risks of sharing personal data with a third party) (Smith et al. 2011), also in the context of genetic data sharing (e.g., Samad et al. 2023; Sanderson et al. 2016; Weidman et al. 2019). In doing so, literature has treated privacy-utility trade-offs as a personal phenomenon that primarily concerns the sharing individuals' personal privacy risks and personal benefits (Bélanger and James 2020), largely neglecting interdependent privacy risks and benefits (see Table A.1 in the appendix for a definition of personal and interdependent privacy risks and benefits). As such, interdependent privacy-utility trade-offs (i.e., cognitive trade-offs made between the benefits and privacy risks for someone else when sharing personal data with a third party) remain a nascent stream of research. In terms of genetic data sharing, previous research has also focused primarily on personal privacy-utility trade-offs. Despite initial research on interdependent privacy considerations in genetic data sharing decisions (e.g., Weidman et al. 2019), previous research has largely neglected the possible existence of interdependent privacy-utility trade-offs or has only broadly operationalized risk and benefit perceptions as higher-order concepts that encompasses both personal and interdependent risks and benefits (Oliver et al. 2012; Sanderson et al. 2017). Thus, there remains a lack of comprehensive understanding of how interdependent privacy-utility trade-offs affect the sharing intentions of genetic data (Weidman et al. 2019). Especially in comparison to personal privacy-utility trade-offs.

Furthermore, it is likely that the interdependent privacy-utility trade-offs are not uniform across all types of relationships but differ depending on the disclosing individuals' affectivity and sympathy toward potentially affected others (i.e., their social distance (SD) to each other). Extant research shows, for example, that individuals feel more obligated to help close family members than distant family members (McManus et al. 2021). Likewise, genetic privacy risks are higher for close blood relatives of disclosing individuals than for more distant blood relatives because they share larger portions of their DNA (Humbert et al. 2013). Yet, previous research on interdependent privacy-utility trade-offs has not yet sufficiently taken this into account and has neglected how SD might affect the perceptions of interdependent privacy-utility trade-offs and their impact on individuals' willingness to share (WTS) their (genetic) data.

Against the background of these limitations in existing research, we therefore examine the following research questions (RQs) in this study:

> *RQ1: How do personal and interdependent privacy-utility trade-offs impact individuals' WTS their genetic data?*
>
> *RQ2: How does SD moderate the effects of interdependent privacy-utility trade-offs on individuals' WTS their genetic data?*

To answer our research questions, we leverage the privacy calculus as our theoretical lens and augment it with the concept of interdependent privacy. Accordingly, we hypothesize the effects of personal privacy risks and benefits as well as interdependent privacy risks and benefits on individuals' WTS their genetic data. Furthermore, drawing on construal level theory (Trope and Liberman 2010), we hypothesize that SD plays a moderating role that either amplifies or reduces the effects of perceived interdependent privacy risks and benefits. To test our hypotheses, we conduct a scenario-based factorial survey in the context of sharing genetic data for a research project. Although our results support the hypothesized direct effects of personal and interdependent privacy risks and benefits, the role of SD as a moderator remains inconclusive. Our study makes three key contributions to the literature. First, we extend the privacy calculus to explicitly incorporate interdependent privacy risks and benefits as a key driver of individuals' genetic data disclosure decisions. Second, based on construal level theory, we introduce the moderating effect of SD in these interdependent privacy-utility trade-offs. Third, for the genetic privacy literature, we answer calls for an explanatory model of genetic data sharing decisions by contextualizing the privacy calculus.

Theoretical Background

Interdependent Privacy

Privacy is a multidimensional concept that has been widely researched and addressed from various perspectives. In IS research, privacy is generally viewed from the perspective of information privacy and is usually understood as an individual's ability to control the dissemination of their personal information (Smith et al. 2011). Although, on the surface, data about an individual may seem to only affect that individual, in our increasingly digitized and interconnected world the information extracted from one's personal data often also exhibits interdependencies (Humbert et al. 2020). Thus, an individual's ability to manage and control the dissemination of their personal information increasingly depends not only on their own self-disclosure behaviors, but also on the privacy behaviors of others (Humbert et al. 2020).

Interdependence of personal information can result from several sources, ranging from mundane behaviors like sharing the photo of multiple individuals online (Henne et al. 2013), to more complex correlations such as individuals liking similar social media postings (Jernigan and Mistree 2009; Lindamood et al. 2009). One common source of personal information interdependence is homophily, which describes the phenomenon that individuals often bond with other individuals of similar gender, age, social class, and preferences (McPherson et al. 2001). Using this information, a multitude of inferences can be made about individuals, including sensitive information like political orientation (Lindamood et al. 2009) or sexual orientation (Jernigan and Mistree 2009). Capitalizing on homophily, past research has investigated the concept of interdependent privacy especially within social networking contexts (Humbert et al. 2020). Another important source of personal information interdependence stems from the kinship property of genetic data. Because the nucleotide pairs that make up the genetic sequence consist of nucleotides from the mother and the father of an individual, information collected through one's genetic sequence can be correlated to characteristics of the parents and other extended members of the family. This enables the precise inference of highly sensitive information such as appearance, age, or disease predispositions (Erlich et al. 2018). Since research has shown that even with limited knowledge about relatives of an individual sharing their genetic data conclusions about said relatives can be drawn (Humbert et al. 2020), it is of utmost importance to consider interdependent privacy risks when researching genetic privacy.

Notions of interdependent privacy have emerged in a variety of scholarly disciplines, mostly in isolation from each other and under different terminology like group-level privacy, collective privacy, multi-party privacy, networked privacy, peer privacy, or multi-subject privacy. We refer interested readers to a review by Humbert et al. (2020) for a more in-depth comparison of these different terms and instead align our own terminology with that of interdependent privacy to better reflect the complex network of privacy risks

and benefits considered in this work. In IS research, there have been calls for more research beyond individual-level privacy for more than a decade (Smith et al. 2011). Yet, it seems that only recently there has been some uptake in research on interdependent privacy. To date, most IS research in that area has focused on the potential negative consequences of someone else disclosing information that an individual has previously shared with them (Lin and Armstrong 2019; Ozdemir et al. 2017). Additionally, Bélanger and James (2020) develop a theoretical model for multi-level privacy that incorporates an interdependent privacy perspective. Only few studies have explicitly conceptualized the role of individuals' concerns for another party's privacy risks (and, for that matter, also their considerations for another party's benefits) in disclosure decisions (Franz and Benlian 2022; Wirth et al. 2019).

Privacy-Utility Trade-Offs in Genetic Data Sharing

Genetic data are digital representations of the whole or parts of the genetic sequence of individuals. They possess several characterizing properties like encoding information about an individual's health and behavior, or being relatively static over the lifetime of an individual (thus potentially allowing for individuals to be traced over long periods) (Naveed et al. 2015). In particular, genetic data serve as both, a unique identifier of individuals that is virtually impossible to fully anonymize, and as a source of kinship information between blood relatives (Naveed et al. 2015). Due to rapid advances in genetics and accompanying technologies (e.g., genetic sequencing), genetic data have become an invaluable tool for various medical and non-medical use cases (Naveed et al. 2015). Despite the many potential benefits of analyzing genetic data, its proliferation into all walks of life also bears tremendous opportunities to infringe on individuals' privacy. Given the immense potential privacy risks involved with sharing genetic data, extant research has extensively studied and discussed privacy risks in genetic data sharing from diverse perspectives, including genetic privacy attacks and protective technology (Naveed et al. 2015), ethical, legal, and social implications of genetic data sharing (Vos et al. 2017), and individuals' WTS their genetic data (Bearth and Siegrist 2020; Shabani et al. 2014). However, although numerous qualitative and quantitative exploratory studies have investigated individuals' motivations to share their genetic data, only few studies have thus far sought to uncover and explain causal relationships between perceived benefits and perceived risks within such privacy-utility trade-offs. See Table 1 for an overview of related studies.

Study	**Summary**	**Benefits**		**Privacy risks**	
		P	**I**	**P**	**I**
Anderson and Agarwal (2009)	The trade-off between individuals' altruism and self-interest in the form of monetary and non-monetary incentives.	X	O	O	
Anderson and Agarwal (2011)	Impact of situation-specific risk factors on the cost-benefit trade-off in healthcare digitization. Genetic data as one type of data.	O	O	X	
Bearth and Siegrist (2020)	The impact of genetic literacy and trust on privacy-utility trade-offs in genetic data sharing.	X	X	X	
Heath et al. (2016)	The trade-off between benefit awareness and privacy concerns regarding genetic data.	X		X	
Kim et al. (2016)	The impact of online discussions about genetic data on privacy-utility trade-offs.	O	O	O	
Oliver et al. (2012)	Perceived benefits and risks for different release options of genetic data in consent forms.	O	X	X	
Samad et al. (2023)	The impact of privacy-utility trade-offs in genetic data sharing versus financial data sharing.	X	O	X	
Sanderson et al. (2017)	The impact of different consent models on privacy-utility trade-offs in biobanking.	X	X	X	
Weidman et al. (2019)	The impact of personal characteristics and situation-specific risk factors on privacy-utility trade-offs in DTC genetic testing.	O		X	X

P = personal; I = interdependent
X = focus of the study; O = implicitly or explicitly part of the study but not in focus

Table 1. Overview of Studies on Privacy-Utility Trade-Offs in Genetic Data Sharing

Overall, we still possess a limited understanding of how individuals' decisions to share (or not share) their genetic data are possibly impacted by their perceived risks and perceived benefits (Bearth and Siegrist 2020; Weidman et al. 2019). Past explanatory research in this area has especially focused on the impact of

personal risks and benefits (e.g., Bearth and Siegrist 2020) or (sometimes implicitly) considered perceived personal risks (PR) and benefits (PB) and perceived interdependent risks (IR) and benefits (IB) as part of the more general concept of perceived risks and benefits. Despite the interdependent nature of genetic data and first valuable investigations into interdependent privacy considerations in the context of genetic data sharing (i.e., Weidman et al. 2019), it thus remains largely unclear whether and how individuals consider IRs and IBs in their privacy-utility trade-offs, whether these effects differ based on one's SD with potentially affected individuals, and how such interdependent privacy-utility trade-offs compare to personal privacy-utility trade-offs.

Research Model and Hypotheses

Building on prior research efforts in understanding individuals' privacy-utility trade-offs (regarding genetic data sharing), we employ the privacy calculus as our theoretical lens and augment it with the concept of interdependent privacy to better understand how interdependent privacy considerations impact individuals' privacy-utility trade-offs in the genetic data sharing decision-making process and to untangle possible effects of perceived PBs, PRs, IBs, and IRs. Figure 1 depicts our research model. Table A.1 in the appendix provides definitions for all key constructs in our model.

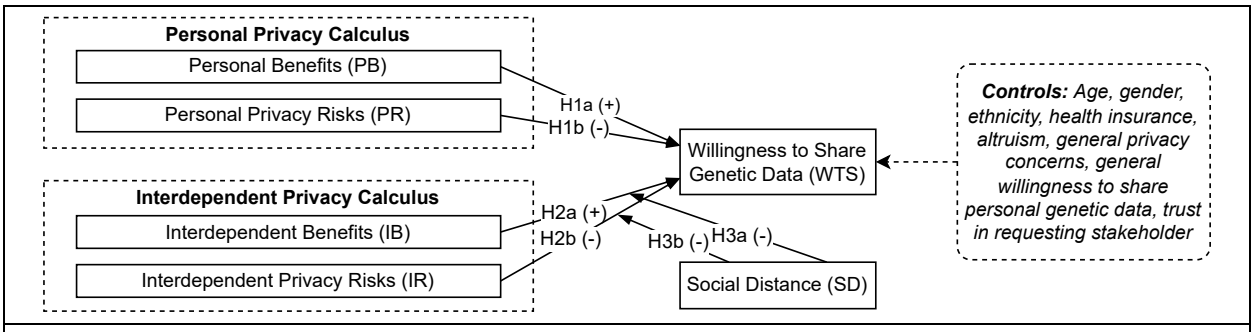

Figure 1. Research Model

Effect of Personal Privacy Calculus on Willingness to Share

Given the possible health benefits of genetic data analyses on the one hand, and the various ways to infringe on individuals' privacy by analyzing their genetic data on the other hand, genetic data sharing is often characterized as a high-risk/high-reward scenario (Heath et al. 2016). Toward that end, past research has extensively investigated individuals' reasons for or against genetic data sharing, highlighting a complex network of motivations (Sanderson et al. 2016; Thiebes et al. 2017). Frequently cited personal motivations for undergoing genetic testing and sharing genetic data with third parties include receiving information about potential disease predispositions (Smith et al. 2007), improved diagnoses and tailored treatments (Vos et al. 2017), or insights about one's genealogical history (Roberts et al. 2017). Likewise, often-named personal motivations against genetic data sharing include the fear of medical implications for oneself (Middleton et al. 2016), concerns about genetic discrimination and stigmatization (Green and Botkin 2003), the commercialization of one's genetic data (Tavani 2004), or the use of one's genetic data into other jurisdictions (Kennett 2019). Extant research especially indicates that individuals often desire some degree of control over the dissemination of their personal genetic data (Jamal et al. 2014; Shabani et al. 2014) and that they make deliberate, situation-specific privacy-utility trade-offs when faced with the decision whether or not to share their genetic data with third parties for a stated purpose (Anderson and Agarwal 2011; Oliver et al. 2012; Shabani et al. 2014). Bearth and Siegrist (2020), for example, found that perceived PBs have a significant positive impact on individuals' WTS their genetic data for research, whereas perceived PRs exhibit a significant negative impact on individuals' WTS their genetic data for research. In line with these empirical insights on individuals' WTS their genetic data we therefore hypothesize that:

> *H1a: High levels of PBs of genetic data sharing are positively associated with an individual's WTS their genetic data with a third party for a stated purpose.*

H1b: High levels of PRs of genetic data sharing are negatively associated with an individual's WTS their genetic data with a third party for a stated purpose.

Effects of Interdependent Privacy Calculus on Willingness to Share

Surveys on individuals' motivations to participate in biobanking (Lemke et al. 2010), direct-to-consumer genetic testing (Goldsmith et al. 2012), or clinical genetic testing (Issa et al. 2009) repeatedly show that next to PBs and PRs, individuals often also explicitly name potential benefits for their relatives (e.g., preventively improving the health of close others through genetic testing) or the wider public (e.g., contributing to medical research on a certain disease), as well as concerns about negative implications for their relatives (e.g., uncertainty about future privacy implications for one's children or upsetting family members) as reasons for, respectively, against genetic data sharing (Gilbar and Barnoy 2012; Green and Botkin 2003). Diergaarde et al. (2007), for example, report that participants in their focus group study understood potential tensions between protecting (their own) privacy and the sharing of genetic data for the benefit of others. Likewise, in their study on individuals' interdependent privacy considerations, Weidman et al. (2019) provide first evidence that concerns for another person's privacy in the context of genetic data sharing might have a negative impact on individuals' WTS their genetic data with DTC genetic testing services.

At the same time, despite the inherent interdependence of genetic data, past explanatory studies have primarily focused on personal privacy-utility trade-offs in genetic data sharing and often only implicitly considered IBs and IRs as dimensions in more general, multi-dimensional benefits and risks constructs. In their study on the effects of privacy literacy and trust on privacy-utility trade-offs in genetic data sharing decisions, for example, Bearth and Siegrist (2020) include several items like "[t]he data are used to advance medical research" or "[t]he data are used to improve public health" in their measurement scale of perceived benefits. Similarly, Sanderson et al. (2017) include items relating to IBs (e.g., "I would feel that taking part could help my family") next to items relating to PBs (e.g., "I would feel that taking part could help me personally") in their measurement of the perceived benefits of participating in biobanking, essentially black-boxing potential differences in the perceptions of personal and interdependent benefits and risks.

In summary, individuals are not only aware of their PBs and PRs risks but also of the IBs and IRs associated with the disclosure of their own genetic data to third parties and there exists first empirical evidence that such aspects play a role in individuals' disclosure decisions. Accordingly, we augment the personal privacy calculus with an interdependent privacy calculus to unblackbox the effects of personal privacy-utility trade-offs versus interdependent privacy-utility trade-offs, and hypothesize that:

H2a: High levels of IBs for another individual are positively associated with an individual's WTS their genetic data with a third party for a stated purpose.

H2b: High levels of IRs for another individual are negatively associated with an individual's WTS their genetic data with a third party for a stated purpose.

Moderating Role of Social Distance on the Interdependent Privacy Calculus

By drawing on construal level theory, we further investigate the effects of SD on the interdependent privacy calculus. According to construal level theory, individuals tend to view events (e.g., privacy infringements) as abstract or concrete depending on how distant or close they are to the individuals' immediate reality (Trope and Liberman 2010). Thus, the more distant an event is from an individual, the more abstract that event is to the individual (i.e., they mentally construe it at a higher level as opposed to a lower, more concrete level). Within construal level theory the distance of an event from an individual's reality is understood in terms of psychological distance, a multidimensional construct consisting of temporal (i.e., how near or far in the future an event will occur), spatial (i.e., how near or far within the physical world an event occurs), hypothetical (i.e., how likely or unlikely an event is to occur), and SD (i.e., how far or close the (groups of) individuals affected by the event are from each other) (Trope and Liberman 2010).

Although there has been little research examining construal level theory in the context of privacy behavior, some initial studies suggest that diverse dimensions of psychological distance may have a role in individuals' privacy behavior (Bandara et al. 2021; Hallam and Zanella 2017). Hallam and Zanella (2017), for example, in their study of temporal distance in the context of social networks, find that individuals trade off future

privacy infringements with immediate benefits in using social networks. Similarly, in their study, Bandara et al. (2021) find a negative moderating effect of psychological distance on the relationship between privacy concerns and defensive privacy behaviors. That is, the higher the perceived psychological distance, the weaker the effect of privacy concerns on defensive privacy behaviors.

Based on the inherent interdependence of genetic data, we focus specifically on the SD dimension and its potential influence on individuals' privacy-utility trade-offs here. The implications of SD have been studied in a variety of contexts, often outside of construal level theory, and from a variety of understandings. However, within privacy research, SD has been primarily studied as an aspect of psychological distance, if at all. A frequently adopted view of SD is that of affective SD, which is usually based on the seminal works by Bogardus (1933) and describes the degree of sympathy an individual has toward another individual. To that end, research building on construal level theory indicates that altruistic behavior toward closer individuals is more morally justifiable than behavior toward more distant individuals (Law et al. 2022). Likewise, individuals might feel more obligated to help close family members than more distant family members (McManus et al. 2021). Regarding IR, risks are higher for close blood relatives of disclosing individuals than for distant blood relatives (Humbert et al. 2013), and parents frequently report concerns over the privacy implications for their children in genetic data sharing situations (Stoffel et al. 2008).

In line with construal level theory, we therefore argue that the closer an individual feels to another individual (e.g., close family members vs. extended family members), the more concretely they will think about the implications of their own genetic data disclosure for that individual. Accordingly, we hypothesize:

> *H3a: SD moderates the effects between IB and WTS such that the (positive) effect of IB on WTS decreases with increasing SD.*
>
> *H3b: SD moderates the effects between IR and WTS such that the (negative) effect of IR on WTS decreases with increasing SD.*

Methods

Factorial Survey Method

To test our hypotheses, we applied the scenario-based factorial survey method, which has increasingly found its way into the IS discipline in recent years (e.g., Lowry 2017; Vance et al. 2015; Willison 2018). The factorial survey method uses contextualized vignettes to enable participants to imagine themselves in the context of a realistic scenario and the scenario's fictional character (Willison 2018). Variations of the vignette represent the different manipulations of the independent variables of the research model (Lowry 2017). After being presented with a vignette, participants are then asked to respond on rating scales that measure the dependent variable. By combining the strengths of experiments and traditional surveys, the factorial survey method enables researchers to manipulate a large number of factors without suffering from multicollinearity problems (Vance et al. 2015). Given the complexity of our theoretical model, the factorial survey method was thus a reasonable choice. Moreover, the vignettes of a factorial survey method can be distributed to different participants, reducing the time and cognitive burden per person. An additional strength of the method is that it is particularly useful in situations where social desirability bias is to be expected since it is based on participants' rating regarding somebody else's hypothetical behavior in a realistic but fictional scenario (Vance et al. 2015). Thereby, individuals are more likely to disclose their own intentions of conducting unethical behavior (Willison 2018). Again, given the sensitive context of our study (i.e., sharing of genetic data with a clinical research project), the factorial survey method seemed like a good choice to test our theoretical model.

As is the case for traditional experiments, factorial surveys are designed based on factors of theoretical interest (i.e., dimensions) (Vance et al. 2015). Each dimension comprises different levels, which build the foundation for constructing the vignette variations. In our case, the research model yields five factors of theoretical interest (i.e., PB, PR, IB, IR, SD) with two levels each (i.e., *low* and *high*). Consequently, our study comprises a full factorial design with 32 (2 to the power of 5) vignettes. As not to overburden participants with too many vignettes, we chose to present only four out of the 32 possible vignettes per participant. When showing participants only a subset of vignettes, it is important to balance and block the vignettes in such a way that all instances of a manipulation are still present once in the block for each participant, also called orthogonal block design. In cases where this is not possible (e.g., due to too many

factors or levels for the number of vignettes shown) the estimation of higher order interaction effects may be limited (Baguley et al. 2022). We utilized the skpr package version 1.0.0 in RStudio version 4.1.1 to find an efficient blocking of four vignettes in eight blocks (4*8 = 32). Optimality inspection showed that the resulting design is D = 1, A = 1, and G = 1 optimal with an I-optimality of 2 (i.e., the block design is fully orthogonal), thus theoretically allowing for analysis of all higher order interaction effects.

Scenario and Vignette Design

For our vignette design we aimed to create a realistic scenario in which a fictional individual named 'Jamie' has to evaluate benefits and risks of sharing their genetic data for themselves as well as for their blood relative named 'Ashley'. The scenario introduces Jamie, who is interested in genetic testing, in a situation where Ashley was diagnosed with brainstem glioma. While searching online about this disease, they find that their local university hospital conducts a research project on brainstem glioma and is currently looking for participants willing to contribute their genetic data. With the scenario in place, we designed the vignettes as email responses in which a scientist from the university hospital answers Jamie's questions regarding the research project. Figure 2 shows two vignettes with all *low* (left) and all *high* (right) treatments.

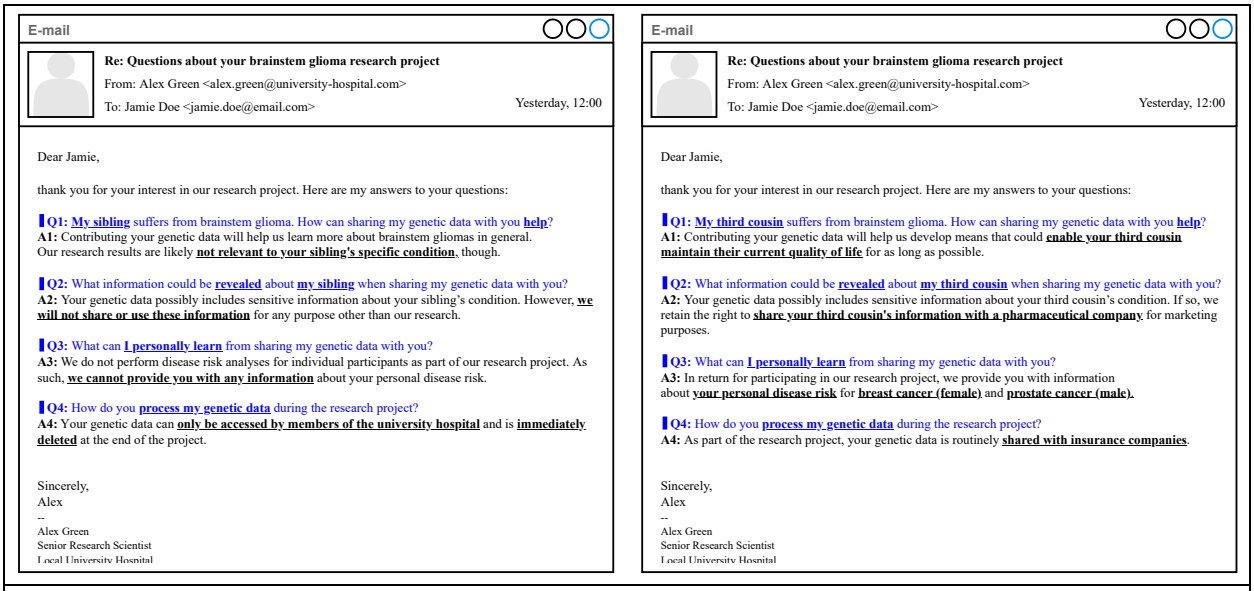

Figure 2. Exemplary Vignettes with *low* Treatments (Left) and *high* Treatments (Right)

To determine the specific high and low manipulations, we first generated five different possible treatments from real world applications and literature for each construct with a hypothesized direct effect on WTS (i.e., PB, PR, IB, IR). Subsequently, we surveyed 48 participants on Amazon Mechanical Turk (mTurk) to deduct a direct and indirect ranking of the treatments from *low* to *high*. We decided to use the lowest and highest ranked treatments as our manipulations in the final vignette design. Furthermore, we decided to construct Ashley as being Jamie's sibling for the *low* SD treatment and Jamie's third cousin for the *high* SD treatment. Additional details on the scenario and vignette development are available in the online supplementary material.[1]

Data Collection and Analysis

With vignettes for the factorial survey in place, we next constructed online survey instruments using soscisurvey.org version 3.2.46. Figure 3 shows an overview of the procedure for the online survey.

The survey first provided a short introduction to the study context (i.e., clinical genetic testing), followed by an initial attention check. Next, we collected pre-vignette measurements about participants personal characteristics, including general WTS genetic data in the respective study context. Thereafter, we

[1] Supplementary material: https://osf.io/ghkcp?view_only=8a58e2d69dbd433baa44087daab35702

presented a short scenario description including a visual representation of Jamie's initial contact (the beginning of Jamie's email to the university hospital). Next, each participant was randomly assigned one of the eight vignette blocks. Randomly choosing the order of the four vignettes in each block, after each presented vignette, participants were then asked to indicate Jamie's WTS their genetic data. To validate whether our treatments worked as expected, we also included manipulation checks for each independent variable. When possible, we adapted measures from prior studies and used multi-item scales to improve reliability and validity of the manipulation check scales. All items were measured using a 7-point Likert scale. A full list of the adapted items can be found in the online supplementary material.[1]

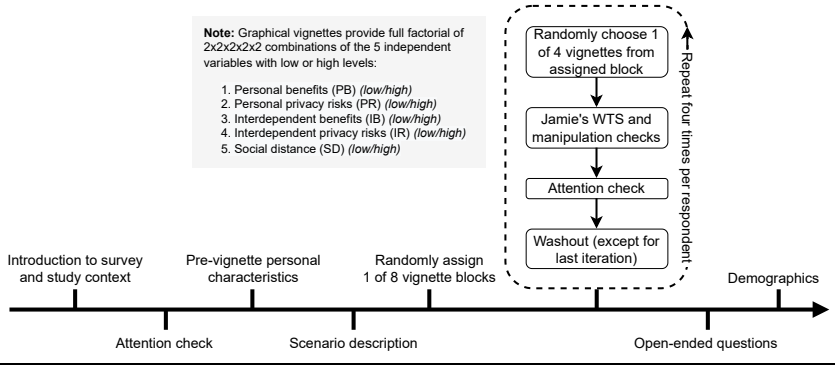

Figure 3. Overview of the Survey Procedure

For the SD manipulation check we used the Inclusion of Other in Self scale (Aron et al. 1992), which is an established visual, single-item scale that has previously been used to measure SD (Hong et al. 2017). Throughout the survey procedure, we used attention checks (i.e., asking participants to select the social distance manipulation presented during a vignette) and washouts (e.g., simple arithmetic questions) between vignettes. The survey closed with open-ended questions regarding the scenario and the collection of selected demographics. Before collecting data, we first conducted a feature test with 9 fellow researchers. Subsequently, we conducted pilot tests with 143 participants from Access by Cint. While the feature and pilot tests lead to minor text and appearance changes, they did not lead to alterations in the general design, as analysis of the pilot dataset indicated functionality of the vignette designs. To determine the planned number of participants for data collection, we calculated the minimal required number of responses per vignette. Following the procedure of Döring and Bortz (2015) for one degree of freedom and a desired significance level of .05, we calculated a minimum of 25.5 responses per vignette. Rounding up to 26 answers for 32 vignettes resulted in a total of 816. Consequently, since each participant was presented with four vignettes, the minimal required number of participants was 204. To ensure sufficient leeway and highest data quality, we opted for a sample size of 600 participants or 2,400 vignette responses.

Because participants were asked to assess multiple vignettes, observations were not independent and unobserved differences in the participants could have introduced a fixed individual effect that biased the vignette ratings for each respondent (Willison 2018). Because ordinary least squares analysis does not handle mixed models of fixed and random effects (Vance et al. 2015), we decided to use a generalized linear mixed model (GLMM). Since the four vignette answers of a single participant are not independent of each other, we followed the recommendation to model this as a repeated measure (Baguley et al. 2022). Furthermore, inspection of the dependent variable (i.e., WTS) indicated no normal distribution but a clear skewness toward larger values, thus requiring Gamma regression. In accordance with our hypothesis model the GLMM contained the main effects for PB, PR, IB, IR and the 2-way interactions for SD*IB and SD*IR as fixed effects. All analyses were conducted using IBM SPSS Statistics v27.

Results

Sample

The survey data were collected from a sample of the general US population. A total of 1,465 people responded to the survey invitation sent out via the panel provider Access by Cint, which translates to a

response rate of about 50.28%. Of those, 662 participants did not finish the survey either because they canceled voluntarily or because they failed the attention checks. Moreover, 191 responses were excluded due to speeding the survey and 142 participants were excluded due to failing our quality checks (i.e., they straight lined, or used certain answering patterns). In sum, 470 valid responses with an average completion time of 24:44 minutes (min: 11:02 minutes; max: 59:55 minutes) remained after data cleansing. When analyzing the 470 valid responses, we recognized that the number of observations were not evenly distributed across groups with group 5 having the lowest number of observations (n=50). As a central requirement of GLMMs is that observations are evenly distributed across the eight treatment groups (Baguley et al. 2022) we decided to randomly remove responses from all groups except group 5 until each group comprised exactly 50 valid responses. Consequently, we were left with 400 valid responses and 1.600 observations (i.e., four answered vignettes per valid response), which is still well above the estimated required sample size of 204 responses (i.e., 816 observations). We summarize important demographics for the remaining sample in Table A.2 appendix. Using a Chi-square test, we assessed all groups to determine whether significant differences exist with respect to important demographics. The results of the Chi-square test indicate that at the 95% confidence interval there were no significant differences between the groups in any demographic dimension.

Treatments Validation

To test the reliability and validity of the measurement model of the included manipulation checks and the dependent variable, we performed exploratory factor analysis (EFA) using IBM SPSS Statistics v27 and confirmatory factor analysis (CFA) using IBM SPSS AMOS v27 for all adapted multi-item scales, including for the adapted WTS scale of our dependent variable. Results of the EFA and CFA led to the exclusion of all reverse coded items for the manipulation checks (i.e., PB5, PR5, IB5, IR5), as well as the removal of one item each for PB (i.e., PB3) and for IR (i.e., IR2). Nevertheless, all constructs (except SD) were still measured with at least three items. In summary, the adjusted measurement model with removed items exhibits adequate levels of convergent validity, discriminant validity, and reliability. With the measurement model in place, we determined the validity of our treatments by a one-way ANOVA (see Table 2).

Construct	**Overall mean**	***low* Treatment**	***high* Treatment**	**ANOVA**	
	N = 1,600	n = 800	n = 800	Effect size	p-Value
PB	4.15 (1.950)	3.16 (1.878)	5.13 (1.465)	546.913	< 0.001
PR	3.29 (1.752)	2.81 (1.561)	3.77 (1.800)	129.265	< 0.001
IB	4.47 (1.901)	3.64 (1.882)	5.31 (1.512)	383.439	< 0.001
IR	3.05 (1.652)	2.75 (1.567)	3.39 (1.732)	59.221	< 0.001
SD[a]	3.90 (2.061)	5.00 (1.817)	2.80 (1.667)	636.939	< 0.001
[a]Higher mean value indicates lower social distance.					

Table 2. Summary of Manipulation Validation

Main Results

Table 3 lists the main results of the analysis of the vignette experiment data.

Category	**Constr.**	**Coefficient**	**Sig.**	**F-value**	**Hypothesis**	
Corrected model		+1.510***	< 0.001	55.770		
Personal calculus	PB	+0.164***	< 0.001	78.389	H1a	Supported
	PR	-0.182***	< 0.001	142.178	H1b	Supported
Interdependent calculus	IB	+0.182***	< 0.001	83.745	H2a	Supported
	IR	-0.100***	< 0.001	48.409	H2b	Supported
Moderation	SD	+0.033*	0.015	5.879		
	IB*SD	-0.087**	0.007	7.200	H3a	Supported
	IR*SD	-0.051†	0.084	2.987	H3b	Not supported
†: p < 0.10; * p < 0.05; ** p < 0.01; *** p < 0.001; Note: SD reported as basis for the interactions.						

Table 3. Fixed Effects Summary

In our model, the base levels were set to match the low treatment conditions. Thus, the coefficients reported for the high treatments indicate the increase or decrease in WTS between the *low* and the *high* treatments. The results show that participants who viewed a vignette with a *high* PB treatment rated significantly higher values of WTS compared to the *low* PB treatment (low, M = 4.243; high, M = 4.999; $p < 0.001$), supporting H1a. Furthermore, participants rated significantly lower WTS after a vignette with *high* PR treatment compared to *low* PR treatment (low, M = 5.044; high, M = 4.205; $p < 0.001$), supporting H1b. Like the personal constructs, the results of the interdependent main effects also supported our hypotheses. Vignettes with the *high* IB treatment resulted in significant higher levels of WTS than the *low* IB treatment (low, M = 4.282; high, M = 4.953; $p < 0.001$), supporting H2a. The *high* IR treatment resulted in significant lower levels of WTS than the *low* IR treatment (low, M = 4.904; high, M = 4.325; $p < 0.001$), supporting H2b.

To test for the hypothesized moderations of SD on the relationships between IB and IR on WTS, we calculated the corresponding two-way interactions. The results are shown in Table 3 and Figure 4. We found a significant interaction effect between SD and IB ($p = 0.007$), supporting H3a as the WTS shows a significantly greater increase from *low* IB (M = 4.266) to *high* IB (M = 5.155) in the case of a *low* SD treatment vignette compared to the increase from *low* IB (M = 4.297) to *high* IB (M = 4.760) for *high* SD treatment vignettes. Regarding the interaction between SD and IR, our results showed no significant effect ($p = 0.084$), indicating no support for H3b. WTS decreased from *low* IR (M = 4.930) to *high* IR (M = 4.461) in the case of a low SD treatment vignette and decreased from *low* IR (M = 4.877) to *high* IR (M = 4.194) for high SD treatment vignettes.

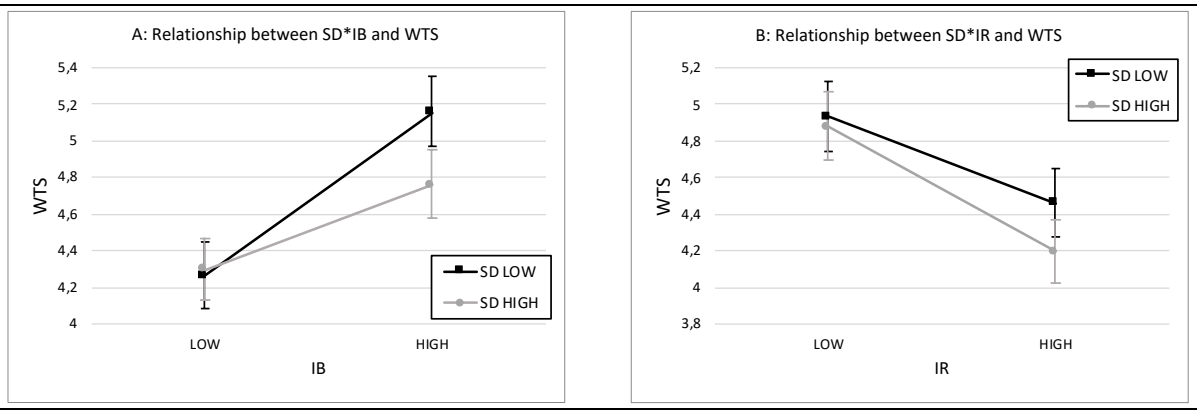

Figure 4. Interaction Graphs for (A) SD*IB Interaction and (B) SD*IR Interaction

In addition to the main GLMM analysis, we augmented the model with 15 collected control variables (i.e., demographics, personal characteristics, order of presented vignettes) that may have confounding effects. We then compared the hypothesis and control model in terms of goodness of fit with the Akaike Information Criterion corrected (AICc) and Bayesian Information Criterion (BIC). For both experiments the hypothesis model showed a greater fit than the control model (cf. Table 4). The addition of control values did not lead to significant changes for the constructs of the hypothesis model.

	Hypothesis model	**Control model**
-2 log likelihood	1258.925	1334.293
AICc	1326.365	1401.770
BIC	1502.226	1576.775

Table 4. Model Quality Criteria

Discussion

Principal Findings

In this study, we developed and empirically tested a research model that explores individuals' decision making with regard to genetic data sharing. The model was built using the privacy calculus as a theoretical

lens and augmented with the concept of independent privacy and elements of construal level theory. The motivation for this focus was based on the observation that extant research has put only little attention on the circumstance that individuals' privacy behavior in genetic data sharing is not just a matter of weighing up personal implications, but inherently includes considerations of implications that one's own genetic data sharing behavior has for others. Toward that end, the further analysis of the results of our experiment provides interesting insights into the model effects that help us to understand better the privacy-utility trade-offs in genetic data sharing and to uncover possible differences between the personal and the interdependent calculus.

First, we find that all direct effects of the personal and interdependent calculus are highly significant (providing support for H1a, H1b, H2a and H2b), substantiating existing notions that individuals make deliberate privacy-utility trade-offs when faced with genetic data disclosure decisions (Samad et al. 2023; Sanderson et al. 2016; Weidman et al. 2019). However, when comparing the coefficients of the direct effects of the personal calculus with those of the interdependent calculus, some interesting differences become apparent. While in the personal calculus the perceived PRs appear to have a significantly stronger (negative) influence on the WTS than the perceived PBs, the relationship in the interdependent calculus is exactly the opposite. Here, perceived IBs appear to have a stronger (positive) influence on WTS than perceived IRs. While it is conceivable that some of this difference in impact is due to our specific scenario design (i.e., sharing genetic data for research purposes), as this may have favored altruistic traits (Anderson and Agarwal 2009), we are nevertheless convinced that these results highlight the existence of a complex web of personal and interdependent privacy-utility trade-offs in genetic data sharing.

Second, our results also provide some interesting insights into the potential moderating role of SD in the impact of the interdependent calculus on individuals' WTS. For example, we found support for H3a (i.e., the closer a disclosing individual feels to a potentially affected other person, the more important IBs become for a disclosure decision), despite the fact that both the *high* and *low* manipulations of SD involved kin (third cousins vs. siblings). This finding is consistent with construal level theory, which posits that the closer an object (i.e., involved kin) is perceived, the more concrete it will be thought of. More specifically, it is in line with contemporary research on construal level theory that shows that altruistic behavior toward closer people is more morally justifiable than behavior toward more distant people (Law et al. 2022). In contrast, we find no support for H3b in our results. Thus, the nature of the disclosing individual's relationship with a potentially affected other does not appear to impact the effect of perceived IRs on WTS, which contradicts construal level theory. We see two possible explanations for this observation. First, the perceived difference between the *low* and *high* condition of IR in our experiment might have simply been too small to tease out the moderating effect of SD. Second, there might be an interaction effect between IB and IR in the *low* SD condition, where IBs overshadow the IRs. If individuals feel that the affected other is very close to them (i.e., *low* SD), they might be willing to accept higher levels of IRs to ensure that they can maintain their quality of life (i.e., *high* IB). Simply put, the survival and quality of life of our closest family members become so important that even severe data breaches become negligible.

Theoretical and Practical Implications

Our study makes three key contributions to the literature. First, we add to the literature on privacy-utility trade-offs in privacy decision-making by extending the privacy calculus and integrating considerations of interdependencies (i.e., IBs and IRs) as important determinants of individuals' disclosure decisions. In doing so, we show that interdependent aspects can play a key role in disclosure decisions and highlight a complex network of PBs and PRs and IBs and IRs. Similar to other studies recently published in major IS outlets (e.g., Bélanger and James 2020), our study thus provides valuable insights toward better understanding information privacy as a group-level phenomenon, which IS scholars have demanded for over a decade (Smith et al. 2011). Second, we add to a nascent stream of privacy literature that investigates privacy behaviors through the lens of construal level theory (e.g., Bandara et al. 2021), by investigating psychological distance at a more granular level. Specifically, by introducing the moderating role of SD to the interdependent calculus, we provide evidence for how a specific subdimension of psychological distance impacts privacy-utility trade-offs. This contrasts with most extant research in that area, which usually conceptualizes psychological distance at a higher level of abstraction and does not explicitly consider its subdimensions in isolation. Third, we add to the literature on genetic privacy and answer calls for more research in that area and help establish a comprehensive explanatory model of genetic data sharing decisions. We do so by integrating three distinct theoretical perspectives (i.e., privacy calculus,

interdependent privacy, and construal level theory) in order to account for the specificities of genetic data, which not only hold sensitive information about oneself, but also about blood relatives that share certain genetic information with a potential donor. For other researchers interested in continuing their research efforts in this area, our study also provides a validated set of vignettes that can be used as treatments for the constructs in the personal and interdependent calculus.

By highlight the importance of both personal and interdependent risks and benefits perceptions with regard to individuals' genetic data sharing decisions, our study also makes some important contributions to practice. For organizations that collect and process genetic data (e.g., clinics, DTC genetic testing services, genetic research labs), our results emphasize that they should keep in mind to not limit themselves to pointing out the PBs that individuals gain from undergoing genetic testing and thereby sharing their genetic data, but also highlight the beneficial information that may be gained for blood-relatives due to the kinship property of genetic data. However, by nature this also leads to associated IRs that arise from sharing genetic data. Extant research has criticized that many providers of genetic tests do not present sufficient information to their customers about how their genetic tests might affect others (Wallace et al. 2015), although this information is highly relevant for individuals in their data sharing decisions, as the results of our study demonstrate. Thus, an obvious step in addressing this issue would be for providers of genetic tests to review their privacy policies with an eye toward transparently illustrating the specific areas in which steps are being taken to reduce both PRs and IRs. Moreover, our results also paint an interesting picture for policy makers. Ever since the early 2000s there have been discussions about a joint account model for genetic information (i.e., whether genetic information belongs to an individual or the whole family and therefore whether disclosure decisions must be made by the whole family) (Foster et al. 2015). Thus far, these discussions have not led to substantive changes in the legal frameworks and data disclosure decisions are still fundamentally a personal decision. Based on the results of our study, a joint account model might not seem all that necessary after all. Given that individuals do already consider personal and interdependent privacy-utility trade-offs when presented with the necessary information, a more pragmatic approach might be to educate individuals about such IRs and IBs in the respective disclosure situations.

Limitations and Future Research

This study is not without limitations. For one, due to the nature of the factorial survey method our study is limited to the examination of how personal and interdependent privacy-utility trade-offs form individuals' behavioral intentions to share their genetic data and does not include investigations of actual behavior or how such intentions translate to actual behavior. While we think that this boundary of our study is an appropriate one, it would be an obvious extension of our work to derive and test a model that also includes genetic data sharing behavior. Moreover, while we aimed to balance data quality and burden for participants some methodological decisions, such as the number of vignettes per participant (in general it is considered optimal to show all possible vignettes), may have affected some of our results. We addressed this by designing our factorial survey study in line with best practices from IS (Lowry 2017; Vance et al. 2015; Willison 2018). For another, we did not consider different genetic data sharing contexts and instead only focused on one specific scenario (i.e., sharing genetic data with a research institution for research on a severe health issue). Genetic data is, however, not only shared for health reasons, but also for other reasons like recreational or legal purposes. It will be an interesting avenue for future research to investigate the different effects of the different benefits and risks provided by genetic data sharing more in depth, especially given the broad diversity of potential application contexts and associated benefits and risks that were not covered by our study. Lastly, research has also started to explore interdependencies of risk and benefits perceptions in the privacy calculus (Kehr et al. 2015; Sun et al. 2021), which we did not consider in our study. We think that this phenomenon deserves additional attention in future research, because it suggests that individuals' privacy-utility trade-offs in genetic data sharing could be even more complex and dynamic than presented here.

Conclusion

The sharing of sensitive data with third parties increasingly poses privacy risks not only to the disclosing individuals, but also to their family and friends. The sharing of genetic data is a prime example of such IRs. With this study, we aimed to understand better how IRs (and IBs) impact individuals' privacy-utility trade-offs in genetic data sharing. To do so, we augmented the privacy calculus as a theoretical lens with an

interdependent perspective and SD as a specific subdimension of construal level theory. We empirically tested the resulting research model by means of using the vignette-based factorial survey method in the context of genetic data sharing for research purposes. While the results clearly show that personal and interdependent privacy-utility trade-offs play an important role in forming individuals' genetic data disclosure decisions, the role of SD remains inconclusive and warrants further investigation. By augmenting the privacy calculus with an interdependent perspective, this study approaches genetic privacy not solely on an individual level of analysis, but rather engages with a decision process that includes both individual-level and group-level considerations simultaneously. In doing so, our research can be of value for scholars in the future, when investigating other information privacy phenomena that include group-level considerations.

References

Anderson, C. L., and Agarwal, R. 2009. "Genetic Information Altruists: How Far and to Whom Does Their Generosity Extend?," in: *Proceedings of the 30th International Conference on Information Systems (ICIS'19),* H. Chen and S. Slaughter (eds.). Phoenix, Arizona, USA: Association for Information Systems, pp. 1-19.

Anderson, C. L., and Agarwal, R. 2011. "The Digitization of Healthcare: Boundary Risks, Emotion, and Consumer Willingness to Disclose Personal Health Information," *Information Systems Research* (22:3), pp. 469-490.

Aron, A., Aron, E. N., and Smollan, D. 1992. "Inclusion of Other in the Self Scale and the Structure of Interpersonal Closeness," *Journal of Personality and Social Psychology* (63:4), pp. 596-612.

Baguley, T., Dunham, G., and Steer, O. 2022. "Statistical Modeling of Vignette Data in Psychology," *British Journal of Psychology* (113), pp. 1143–1163.

Bandara, R. J., Fernando, M., and Akter, S. 2021. "Construing Online Consumers' Information Privacy Decisions: The Impact of Psychological Distance," *Information & Management* (58:7), pp. 1-11.

Bearth, A., and Siegrist, M. 2020. "Psychological Factors That Determine People's Willingness-to-Share Genetic Data for Research," *Clinical Genetics* (97:3), pp. 483-491.

Bélanger, F., and James, T. L. 2020. "A Theory of Multilevel Information Privacy Management for the Digital Era," *Information Systems Research* (31:2), pp. 510-536.

Bogardus, E. S. 1933. "A Social Distance Scale," *Sociology & Social Research* (17), pp. 265-271.

Carballo, R., Schmall, E., and Tumin, R. 2024. "23andme Breach Targeted Jewish and Chinese Customers, Lawsuit Says." Retrieved 2024/04/29, 2024, from https://www.nytimes.com/2024/01/26/business/23andme-hack-data.html

Diergaarde, B., Bowen, D. J., Ludman, E. J., Culver, J. O., Press, N., and Burke, W. 2007. "Genetic Information: Special or Not? Responses from Focus Groups with Members of a Health Maintenance Organization," *American Journal of Medical Genetics Part A* (143:6), pp. 564-569.

Döring, N., and Bortz, J. 2015. *Forschungsmethoden Und Evaluation in Den Sozial- Und Humanwissenschaften*. Berlin Heidelberg: Springer.

Erlich, Y., Shor, T., Pe'er, I., and Carmi, S. 2018. "Identity Inference of Genomic Data Using Long-Range Familial Searches," *Science* (362:6415), pp. 690-694.

Foster, C., Herring, J., and Boyd, M. 2015. "Testing the Limits of the ‚Joint Account' Model of Genetic Information: A Legal Thought Experiment," *Journal of Medical Ethics* (41:5), pp. 379–382.

Franz, A., and Benlian, A. 2022. "Exploring Interdependent Privacy–Empirical Insights into Users' Protection of Others' Privacy on Online Platforms," *Electronic Markets* (32:4), pp. 2293-2309.

Gilbar, R., and Barnoy, S. 2012. "Disclosure of Genetic Information to Relatives in Israel: Between Privacy and Familial Responsibility," *New Genetics and Society* (31:4), pp. 391-407.

Goldsmith, L., Jackson, L., O'connor, A., and Skirton, H. 2012. "Direct-to-Consumer Genomic Testing: Systematic Review of the Literature on User Perspectives," *European Journal of Human Genetics* (20:8), pp. 811-816.

Green, M. J., and Botkin, J. R. 2003. ""Genetic Exceptionalism" in Medicine: Clarifying the Differences between Genetic and Nongenetic Tests," *Annals of Internal Medicine* (138:7), pp. 571-575.

Hallam, C., and Zanella, G. 2017. "Online Self-Disclosure: The Privacy Paradox Explained as a Temporally Discounted Balance between Concerns and Rewards," *Computers in Human Behavior* (68:2017), pp. 217-227.

Heath, D., Ardestani, A., and Nemati, H. 2016. "Sharing Personal Genetic Information: The Impact of Privacy Concern and Awareness of Benefit," *Journal of Information, Communication and Ethics in Society* (14:3), pp. 288-308.

Henne, B., Szongott, C., and Smith, M. 2013. "Snapme If You Can: Privacy Threats of Other Peoples' Geo-Tagged Media and What We Can Do About It," in: *Proceedings of the 6th ACM Conference on Security and Privacy in Wireless and Mobile Networks (WiSec '13),* L. Buttyán, A.-R. Sadeghi and M. Gruteser (eds.). Budapest, Hungary: Association for Computing Machinery, pp. 95–106.

Hong, K., Pavlou, P. A., Shi, N., and Wang, K. 2017. "On the Role of Fairness and Social Distance in Designing Effective Social Referral Systems," *MIS Quarterly* (41:3), pp. 787-809.

Humbert, M., Ayday, E., Hubaux, J.-P., and Telenti, A. 2013. "Addressing the Concerns of the Lacks Family: Quantification of Kin Genomic Privacy," in: *Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS'13),* A.-R. Sadeghi, V. Gligor and M. Yung (eds.). Berlin, Germany: ACM, pp. 1141-1152.

Humbert, M., Trubert, B., and Huguenin, K. 2020. "A Survey on Interdependent Privacy," *ACM Computing Surveys* (52:6), pp. 1-40.

Issa, A. M., Tufail, W., Hutchinson, J., Tenorio, J., and Baliga, M. P. 2009. "Assessing Patient Readiness for the Clinical Adoption of Personalized Medicine," *Public Health Genomics* (12:3), pp. 163-169.

Jamal, L., Sapp, J. C., Lewis, K., Yanes, T., Facio, F. M., Biesecker, L. G., and Biesecker, B. B. 2014. "Research Participants' Attitudes Towards the Confidentiality of Genomic Sequence Information," *European Journal of Human Genetics* (22:8), pp. 964-968.

Jernigan, C., and Mistree, B. F. T. 2009. "Gaydar: Facebook Friendships Expose Sexual Orientation," *First Monday* (14:10).

Kehr, F., Kowatsch, T., Wentzel, D., and 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), pp. 607-635.

Kennett, D. 2019. "Using Genetic Genealogy Databases in Missing Persons Cases and to Develop Suspect Leads in Violent Crimes," *Forensic Science International* (301:2019), pp. 107-117.

Kim, S. C., Cappella, J. N., and Price, V. 2016. "Online Discussion Effects on Intention to Participate in Genetic Research: A Longitudinal Experimental Study," *Psychology & Health* (31:9), pp. 1025-1046.

Law, K. F., Campbell, D., and Gaesser, B. 2022. "Biased Benevolence: The Perceived Morality of Effective Altruism across Social Distance," *Personality and Social Psychology Bulletin* (48:3), pp. 426-444.

Lemke, A., Wolf, W., Hebert-Beirne, J., and Smith, M. 2010. "Public and Biobank Participant Attitudes toward Genetic Research Participation and Data Sharing," *Public Health Genomics* (13:6), pp. 368-377.

Limayem, M., Hirt, S. G., and Cheung, C. M. K. 2007. "How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance," *MIS Quarterly* (31:4), pp. 705-737.

Lin, S., and Armstrong, D. 2019. "Beyond Information: The Role of Territory in Privacy Management Behavior on Social Networking Sites," *Journal of the Association for Information Systems* (20:4), p. 2.

Lindamood, J., Heatherly, R., Kantarcioglu, M., and Thuraisingham, B. 2009. "Inferring Private Information Using Social Network Data," in: *Proceedings of the 18th International Conference on World Wide Web*. Madrid, Spain: Association for Computing Machinery, pp. 1145-1146.

Lowry, P. B., Moody, G. D., & Chatterjee, S. 2017. "Using IT Design to Prevent Cyberbullying," *Journal of Management Information Systems* (34:3), pp. 863-901.

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

McManus, R. M., Mason, J. E., and Young, L. 2021. "Re-Examining the Role of Family Relationships in Structuring Perceived Helping Obligations, and Their Impact on Moral Evaluation," *Journal of Experimental Social Psychology* (96).

McPherson, M., Smith-Lovin, L., and Cook, J. M. 2001. "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology* (27:1), pp. 415-444.

Middleton, A., Milne, R., Thorogood, A., Kleiderman, E., Niemiec, E., Prainsack, B., Farley, L., Bevan, P., Steed, C., Smith, J., Vears, D., Atutornu, J., Howard, H. C., and Morley, K. I. 2019. "Attitudes of Publics Who Are Unwilling to Donate DNA Data for Research," *European Journal of Medical Genetics* (62:5), pp. 316-323.

Middleton, A., Morley, K. I., Bragin, E., Firth, H. V., Hurles, M. E., Wright, C. F., and Parker, M. 2016. "Attitudes of Nearly 7000 Health Professionals, Genomic Researchers and Publics toward the Return of Incidental Results from Sequencing Research," *European Journal of Human Genetics* (24:1), pp. 21-29.

Naveed, M., Ayday, E., Clayton, E. W., Fellay, J., Gunter, C. A., Hubaux, J.-P., Malin, B. A., and Wang, X. 2015. "Privacy in the Genomic Era," *ACM Computing Surveys* (48:1), pp. 1-44.

Oliver, J. M., Slashinski, M. J., Wang, T., Kelly, P. A., Hilsenbeck, S. G., and McGuire, A. L. 2012. "Balancing the Risks and Benefits of Genomic Data Sharing: Genome Research Participants' Perspectives," *Public Health Genomics* (15:2), pp. 106-114.

Ozdemir, Z. D., Jeff Smith, H., and 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), pp. 642-660.

Roberts, J. S., Gornick, M. C., Carere, D. A., Uhlmann, W. R., Ruffin, M. T., and Green, R. C. 2017. "Direct-to-Consumer Genetic Testing: User Motivations, Decision Making, and Perceived Utility of Results," *Public Health Genomics* (20:1), pp. 36-45.

Samad, Z., Wooders, M., Malin, B., and Vorobeychik, Y. 2023. "Risk, Trust, and Altruism in Genetic Data Sharing," *Journal of Public Economic Theory* (25:6), pp. 1251-1269.

Sanderson, S. C., Brothers, K. B., Mercaldo, N. D., Clayton, E. W., ..., and Holm, I. A. 2017. "Public Attitudes toward Consent and Data Sharing in Biobank Research: A Large Multi-Site Experimental Survey in the Us," *The American Journal of Human Genetics* (100:3), pp. 414-427.

Sanderson, S. C., Linderman, M. D., Suckiel, S. A., Diaz, G. A., Zinberg, R. E., Ferryman, K., Wasserstein, M., Kasarskis, A., and Schadt, E. E. 2016. "Motivations, Concerns and Preferences of Personal Genome Sequencing Research Participants: Baseline Findings from the Healthseq Project," *European Journal of Human Genetics* (24:1), pp. 14-20.

Shabani, M., Bezuidenhout, L., and Borry, P. 2014. "Attitudes of Research Participants and the General Public Towards Genomic Data Sharing: A Systematic Literature Review," *Expert Review of Molecular Diagnostics* (14:8), pp. 1053-1065.

Smith, A., Moran, A., Boyd, M. C., Bulman, M., Shenton, A., Smith, L., Iddenden, R., Woodward, E. R., Lalloo, F., Maher, E. R., and Evans, D. G. 2007. "Phenocopies in Brca1 and Brca2 Families: Evidence for Modifier Genes and Implications for Screening," *Journal of Medical Genetics* (44:1), pp. 10-15.

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

Stoffel, E. M., Ford, B., Mercado, R. C., Punglia, D., Kohlmann, W., Conrad, P., Blanco, A., Shannon, K. M., Powell, M., Gruber, S. B., Terdiman, J., Chung, D. C., and Syngal, S. 2008. "Sharing Genetic Test Results in Lynch Syndrome: Communication with Close and Distant Relatives," *Clinical Gastroenterology and Hepatology* (6:3), pp. 333-338.

Sun, Y., Wang, N., and Shen, X.-L. 2021. "Calculus Interdependency, Personality Contingency, and Causal Asymmetry: Toward a Configurational Privacy Calculus Model of Information Disclosure," *Information & Management* (58:8), p. 103556.

Tavani, H. T. 2004. "Genomic Research and Data-Mining Technology: Implications for Personal Privacy and Informed Consent," *Ethics and Information Technology* (6:1), pp. 15-28.

Thiebes, S., Lyytinen, K. J., and Sunyaev, A. 2017. "Sharing Is About Caring? Motivating and Discouraging Factors in Sharing Individual Genomic Data," in: *Proceedings of the 38th International Conference on Information Systems (ICIS'17),* Y.J. Kim, R. Agarwal and J.K. Lee (eds.). Seoul, South Korea: Association for Information Systems, pp. 1-20.

Thiebes, S., Toussaint, P. A., Ju, J., Ahn, J.-H., Lyytinen, K., and Sunyaev, A. 2020. "Valuable Genomes: Taxonomy and Archetypes of Business Models in Direct-to-Consumer Genetic Testing," *Journal of Medical Internet Research* (22:1), p. e14890.

Trope, Y., and Liberman, N. 2010. "Construal-Level Theory of Psychological Distance," *Psychology Review* (117:2), pp. 440-463.

Vance, A., Lowry, P. B., and 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), pp. 345-366.

Vos, S., van Delden, J. J. M., van Diest, P. J., and Bredenoord, A. L. 2017. "Moral Duties of Genomics Researchers: Why Personalized Medicine Requires a Collective Approach," *Trends in Genetics* (33:2), pp. 118-128.

Wallace, S. E., Gourna, E. G., Nikolova, V., and Sheehan, N. A. 2015. "Family Tree and Ancestry Inference: Is There a Need for a 'Generational' Consent?," *BMC Medical Ethics* (16:87), pp. 1-9.

Weidman, J., Aurite, W., and Grossklags, J. 2019. "On Sharing Intentions, and Personal and Interdependent Privacy Considerations for Genetic Data: A Vignette Study," *IEEE/ACM Transactions on Computational Biology and Bioinformatics* (16:4), pp. 1349-1361.

Willison, R., Warkentin, M., & Johnston, A. C. 2018. "Examining Employee Computer Abuse Intentions: Insights from Justice, Deterrence and Neutralization Perspectives.," *Information Systems Journal* (28:2), pp. 266-293.

Wirth, J., Maier, C., Laumer, S., and Weitzel, T. 2019. "Perceived Information Sensitivity and Interdependent Privacy Protection: A Quantitative Study," *Electronic Markets* (29:3), pp. 359-378.

Appendix

Construct *(Abbr.)*	**Adapted definition**	**Based on**
Willingness to share *(WTS)*	An individual's willingness to share their personal genetic data with a third party for a stated purpose.	Anderson and Agarwal (2011)
Personal benefits *(PB)*	The perceived benefits for the sharing individual when sharing their personal genetic data with a third party for a stated purpose.	Limayem et al. (2007)
Personal privacy risks *(PR)*	An individual's expectation of personal losses associated with the disclosure of their personal genetic information to a third party for a stated purpose.	Malhotra et al. (2004)
Interdependent benefits *(IB)*	The benefits pertaining to another individual B as perceived by a sharing individual A, when sharing their personal genetic data with a third party for a stated purpose.	*Self-developed based on personal benefits construct*
Interdependent privacy risks *(IR)*	An individual A's expectation of losses for another affected individual B, associated with individual A disclosing their personal genetic information to a third party for a stated purpose.	*Self-developed based on personal privacy risks construct*
Social distance *(SD)*	The affectivity (i.e., how much sympathy) a sharing individual A feels for another individual B that might be affected by the sharing individual A's decision to disclose their genetic data.	Hong et al. (2017)

Table A.1. Definition of Key Constructs

Demographic		**Total (%)**	**Experimental group**								**Chi-square test**	
			1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**F**	**p-Value**
N	-	400 (100.00%)	50	50	50	50	50	50	50	50	-	-
Gender	Female	236 (59.00%)	30	28	27	35	28	27	30	31	12.095	0.599
	Male	160 (40.00%)	18	22	22	15	21	23	20	19		
	Other	4 (1.00%)	2	0	1	0	1	0	0	0		
Age (in years)	18-24	47 (11.75%)	5	5	5	6	6	9	6	5	28.877	0.966
	25-34	75 (18.75%)	5	9	11	9	10	8	13	10		
	35-44	84 (21.00%)	11	9	8	10	9	10	13	14		
	45-54	69 (17.25%)	9	10	7	9	11	8	8	7		
	55-64	55 (13.75%)	8	7	8	5	8	6	4	9		
	65-74	56 (14.00%)	12	8	8	8	4	7	6	3		
	75+	14 (3.50%)	0	2	3	3	2	2	0	2		
	Prefer not to say	0 (0.00%)	0	0	0	0	0	0	0	0		
Ethnicity	American Indian or Alaska Native	5 (1.25%)	0	1	0	1	0	1	0	2	28.028	0.792
	Asian	17 (4.25%)	3	1	2	0	4	4	2	1		
	Black or African American	45 (11.25%)	5	6	5	7	7	3	5	7		
	Native Hawaiian or Other Pacific Islander	2 (0.50%)	1	0	0	0	1	0	0	0		
	White	324 (81.00%)	41	41	41	41	36	42	43	39		
	Prefer not to say	7 (1.75%)	0	1	2	1	2	0	0	1		
Health	Excellent	35 (8.75%)	3	9	3	5	4	4	3	4	28.884	0.757
	Very good	114 (28.50%)	12	15	13	14	9	19	17	15		
	Good	138 (34.50%)	21	11	21	18	19	18	14	16		
	Fair	93 (23.25%)	13	11	11	11	15	7	14	11		
	Poor	19 (4.75%)	1	3	2	2	3	2	2	4		
	Prefer not to say	1 (0.25%)	0	1	0	0	0	0	0	0		

Table A.2. Summary Panel Demographics