

Attitudes on Data Use for Public Benefit: Investigating Context-Specific Differences Across Germany, Spain, and the United Kingdom With a Longitudinal Survey Experiment

Social Media + Society
October-December 2024: 1–18
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20563051241301202
journals.sagepub.com/home/sms
 Sage

Frederic Gerdon^{1,2}

Abstract

With technological advances, governments and companies gain opportunities to collect data to provide public benefits. However, such data collections and uses need to fulfill ethical standards and comply with citizens' privacy preferences, which may vary across several dimensions. The Comparative Privacy Research Framework suggests specific comparative dimensions that may shape such privacy-related perceptions. I propose how to integrate into this framework a specific meso-level perspective for concisely operationalizing data uses context-specifically: the privacy theory of contextual integrity, developed by Helen Nissenbaum. This article presents an empirical application of this approach by investigating specific data use scenarios across countries, while simultaneously considering temporal, international, and individual-level variation. To this end, an online survey experiment was conducted in three countries (Germany, Spain, and the United Kingdom) in December 2022 and May 2023. In this experiment, respondents rated the appropriateness of fictitious data use scenarios. The scenarios varied by data type, data recipients, and conditions of data use. The results show that the effects of contextual parameters vary across countries to different degrees. Respondents react particularly sensitively to changes in data types, with health data being overall most accepted to be used. The relative acceptance of the data recipients clearly varies across countries. Country-level individualism is not consistently related to the desired level of control over data. These findings highlight the usefulness of contextual integrity to unmask meso-level, context-specific variations in privacy attitudes across countries. A meso-level perspective that operationalizes data uses according to contextual integrity can therefore inform comparative privacy research and privacy-related policymaking.

Keywords

privacy attitudes, survey, survey experiment, contextual integrity, Comparative Privacy Research

Introduction

Numerous data collection practices that aim to provide individual benefits produce data that may be also used for a public benefit. Digital patient records, smartphone movement collections, smart home and smart grid technologies, social media data—all these data can be used to provide immediate benefits to individuals, but also could be used for scholarly research or the improvement of public management. Novel data collection efforts and an internationalization of data markets, such as envisioned by Common European Data Spaces (European Commission, 2023), increase the opportunities for such public benefit data uses. However, these opportunities come with privacy concerns, for example,

voiced by scholars who worry about undue surveillance (Newlands et al., 2020; Vitak & Zimmer, 2020). For ethical data collection and use, we need to design data use practices that are informed by citizens' preferences, among others.

However, one-off privacy surveys on public preferences focusing on specific perceptions at a specific time and place

¹University of Mannheim, Germany

²Ludwig-Maximilians-Universität in Munich, Germany

Corresponding Author:

Frederic Gerdon, University of Mannheim, A5, 6, 68159 Mannheim, Germany.

Email: frederic.gerdon@uni-mannheim.de



are not readily generalizable across countries and contexts. Given the internationalization and cross-sectoral application of data regulations, more fine-grained comparisons become increasingly important. In the present article, I combine the comparative perspective of the Comparative Privacy Research Framework (Masur et al., 2024) with the context-based notion of privacy as “contextual integrity” (Nissenbaum, 2010) to show that investigating countries with respect to general privacy attitudes might miss important nuances in people’s perspectives on which kinds of data uses are acceptable. Contextual integrity can therefore meaningfully enhance the comparative power of the Comparative Privacy Research Framework by providing a template for context-specific comparisons across countries.

To this end, I empirically investigate how attitudes related to data use for public benefit vary across social contexts within and between countries. I conduct an international survey experiment that presents respondents with text descriptions of hypothetical data use scenarios. These scenarios vary by parameters as suggested by contextual integrity (data type, involved actors, conditions of data use), such that effects of changes in parameters on respondents’ acceptance can be estimated. The study was fielded in three countries with different levels of individualism, which was previously shown to be related to the acceptance of public benefit data use (Li et al., 2017; see section “Sample”): Germany, Spain, and the United Kingdom (UK). Moreover, I conduct the study at two time points as privacy attitudes may change over time with the salience of public issues related to a specific data use (Gerdon et al., 2021), and finally also consider individual-level predictors of acceptance.

In summary, the main research question is: *Under which conditions do people deem data use for the public benefit appropriate, and do such attitudes vary across social contexts and countries over time?* I answer this question with respect to the outlined four components of comparison: contextual, international, interindividual, and longitudinal. The resulting evidence on variations of privacy attitudes along these components allows (1) researchers to learn about the variability of privacy attitudes across countries as depending on social context and over time and thereby (2) policymakers to consider people’s preferences for an appropriate regulation of data use for public benefit.

Attitudes on Data Use for Public Benefit: Comparisons by Four Components

The collection of specific pieces of information about individuals may serve different kinds of purposes. For instance, physicians may collect health data of patients to provide diagnoses and treatments. The very same collected health data may also be used by researchers to study, for example, risk factors for specific diseases. The former purpose

provides a direct personal benefit to the individual, while the latter purpose can lead to public benefits (such as better treatment options) that may translate into personal benefits.¹

More generally, following the definition by the National Data Guardian for Health and Social Care in England (2022), a “public benefit” arises from data use if the achieved benefits are not outweighed by risks, while benefits can also be of indirect nature. Furthermore, according to this definition, to be a “public benefit,” the broader public or a subsection of the public need to benefit, such that exclusively commercial benefit does not fall under this definition. In addition, the legitimacy of a public benefit data use hinges on whether it has a “social license” (Carter et al., 2015; Shaw et al., 2020) granted by the population. We therefore need to learn whether and under which conditions the (re-)use of individual data for public benefit purposes is deemed acceptable by the public. With the internationalization of regulations on data collection and use, it becomes increasingly important to also know how populations of different countries differ in their acceptance of specific data uses. Such knowledge could aid in formulating policies by showing how populations might react differently to specific data use endeavors (e.g., differences in likelihoods of opting in or out of sharing data from digital patient records with researchers).

In the present article, I argue that such international comparisons can be substantially enriched by comparing perceptions relating to *specific* data use contexts, additional to rather general privacy perceptions. To this end, I draw on the Comparative Privacy Research Framework (Masur et al., 2024) and extend it with the perspective of “contextual integrity” (Nissenbaum, 2010). The Comparative Privacy Research Framework cautions against the over-generalization of findings from privacy studies that focus on single units and offers a structured approach to, among others, international comparisons (Masur et al., 2024). More precisely, Masur et al. (2024) propose to study privacy-related phenomena, such as data use attitudes or behaviors, by comparing at least two “units of comparison” (on the macro-, meso-, or micro-level). They define five types of “categories” (cultural, social, political, economic, and technological) that may influence phenomena or moderate processes.

Understanding privacy as contextual integrity (as proposed by Nissenbaum, 2010) can enrich the Comparative Privacy Research Framework by drawing attention to the specific configuration of social contexts and their respective privacy norms. Nissenbaum understands social contexts as areas of social life such as healthcare and work that come with specific practices, roles, purposes, and norms (Nissenbaum, 2018). For instance, in the healthcare context, there may be specific rules, practices, and expectations of how data collected about a patient by a physician may be used and shared. Masur et al. (2024) explicitly refer to contextual integrity as providing a context-sensitive perspective on privacy. These meso-level social contexts are embedded in larger structural units such as political systems. What contextual integrity adds to the

Table 1. Units of Comparison That I Simultaneously Compare in This Article, Based on the Comparative Privacy Research Framework (Masur et al., 2024).

Level	Units of comparison	Concepts
Macro	Culture (country-level)	Individualism
Meso	Social context	Contextual integrity parameters: data type, actors, transmission principle
Micro	Individual perceptions and behaviors	E.g., privacy concerns and trust in data recipients (see section “Method”)
<i>Longitudinal comparison</i>		

Comparative Privacy Research Framework is a concrete template to operationalize data uses within meso-level social contexts. Comparisons of attitudes toward data uses in these meso-level contexts can enhance macro-level country comparisons by unmasking context-specific differences in privacy attitudes beyond “general” privacy attitudes in the investigated countries (relatedly, see Martin & Nissenbaum, 2017). At the same time, country comparisons can reveal how the relevance of specific contextual factors varies across countries (e.g., Li et al., 2017).²

Additional to these macro- and meso-level comparisons, on the micro-level, individuals may display different more general stances toward privacy (Gerber et al., 2018). For instance, age and gender, general privacy perceptions (Smith et al., 2011), and altruism (Y. Kim & Stanton, 2016; Silber et al., 2022) may shape how acceptable individuals deem data use for public benefit in a given situation. All these structural units may interact with each other to affect privacy attitudes (Masur et al., 2024), such that the effects of individual characteristics may vary by country and be more relevant in some social contexts than in others.

Finally, norms and attitudes within units of comparison may change over time due to changes in the societal environment, as previous research has demonstrated with respect to a (potentially temporary) increase of acceptance of use of health data for disease containment early in the COVID-19 pandemic (Gordon et al., 2021). The Comparative Privacy Research Framework is explicitly open for longitudinal comparisons, due to the potential of major events to affect privacy perceptions (Masur et al., 2024).

Table 1 summarizes the comparative approach of the present article, which I will further explain in the following sections. In the following, I apply this theoretical background to the concrete case of privacy attitudes toward data use for public benefit.

Contextual Variation

As argued above, country-level macro comparisons of general privacy attitudes may miss important meso-level contextual differences (relatedly, see Martin & Nissenbaum,

2017). The notion of contextual integrity can enhance such comparisons by asserting that the appropriateness of data flows depends on compliance with privacy norms that are specific to the social contexts in which they are embedded (Nissenbaum, 2010). To assess the appropriateness of data flows, contextual integrity requires us to define *which* data are at stake under involvement of *which actors* and *under which conditions*. For instance, individuals may be willing to share detailed health information with doctors. At the same time, they might find it outraging if employers requested the voluntary sharing of such data. To concretely analyze and assess the appropriateness of a data flow, Nissenbaum provides a data flow template that consists of five parameters: data type, data subject, data sender, data recipient, and transmission principles (i.e., the prerequisites under which the data flow occurs).

Previous research has repeatedly shown that individual evaluations of data flows are sensitive toward changes in the specifications of data flow parameters (e.g., Martin & Nissenbaum, 2017; Terpstra et al., 2023; Utz et al., 2021). I now turn to discussing the contextual integrity parameters in more detail with respect to data use for public benefit and develop hypotheses and research questions.

With respect to data types, contextual integrity does not suggest that any data type is *as such* necessarily more sensitive than another, since sensitivity depends on context (Martin & Nissenbaum, 2017). Empirically, relatively much research has been dedicated to the specific case of health data use. For several kinds of health data, literature reviews have identified that public benefit uses are overall acceptable if the data are safe, the recipients are deemed trustworthy, and commercial interests are not the main focus, among others (Aitken et al., 2016; Hutchings et al., 2020; Kalkman et al., 2022). For social media data, research found that the acceptance of research uses depends on factors such as the research purpose, with a preference toward context-specific user experience research (Gilbert et al., 2021), but the acceptance of research may vary across social media platforms (Gilbert et al., 2023).

Acceptance of public benefit data use may further be affected by salient societal issues. One useful theoretical perspective is provided by Büchi et al. (2022) who draw on the theory of planned behavior and argue that privacy-related scandals may, in the long run, lead to more chilled digital communication behavior. I apply this argumentation to other societal events that make specific issues salient and could therefore (temporarily) affect individual’s attitudes on data use for issue-related contexts. Furthermore, I argue that salience may also lead to more *appreciative* attitudes toward data use, depending on the specific salient event. For instance, previous research has demonstrated that the COVID-19 pandemic increased the acceptance for the use of health data collected on smartphones for public benefit (Gordon et al., 2021). However, such effects may be temporary (see below). For example, several attitudes relating to

surveillance for public security were more favorable in the United States right after 9/11, but this attenuated in the following months and years (Best et al., 2006).

To test this relationship, I compare the acceptance of health data use with the acceptance toward other data types that vary in their relatedness to currently debated public issues, such as the COVID-19 pandemic. This includes energy use, data which became a potentially salient issue in 2022 in the face of increasing energy prizes, and the need to save energy. As less immediately salient data types, I investigate the yet important and frequently debated types of location data (e.g., see the critical discussion by Walsh, 2023) and social media data (Proferes & Walker, 2020):

H1: The use of health data for public benefit is more accepted than the use of other data types for public benefit that are less directly related to the pandemic.

With respect to data recipients, relatively much comparative research is available for health data. Studies show that researchers or associated institutions appear as more accepted health data recipients than government agencies, while companies are least accepted (K. K. Kim et al., 2015) and, for Germany, that pharmaceutical companies are less accepted than researchers associated with universities or research-related public agencies (Haug et al., 2023). While these results suggest relatively high acceptance of health data use by public entities, not all studies share this finding (Gerdon et al., 2021; or for public benefit purposes: Deruelle et al., 2023), and the findings may further vary by concrete data recipient and consent procedure. Some studies suggest that individuals could find the private sector using specific types of health data acceptable if public benefits stood above profit (see Aitken et al., 2016).

From a contextual integrity perspective, as I focus on public benefit purposes, a tendency toward higher acceptance for public recipients can be expected. Public recipients usually more frequently take part in contexts that have the explicit goal to foster public welfare than private recipients and therefore be deemed appropriate. Within public recipients, I expect that respondents consider researchers affiliated with public institutions to be the least likely expected to use data for out-of-context purposes and therefore may enjoy the highest acceptance ratings. These relationships may vary by concrete data type and the conditions of data use:

H2: Public actors, and particularly public researchers, are more accepted than private actors as recipients of data to use for public benefit. The effect of recipient interacts with data type and transmission principles.

With respect to transmission principles, several conditions to share data for public benefit exist. In opt-in scenarios, data are only used after the individual explicitly consents to data use. In the context of health data use for research, a review study found opt-in as the most favored approach, while results varied when deidentified data were to be shared

(Stockdale et al., 2019). The review also concludes that individuals may change their opinions upon learning more about the benefits for research. However, review studies found that consent for use of medical records correlates with individual characteristics and that data sets that only include consented data may be biased (De Man et al., 2023; Kho et al., 2009). Opt-out approaches partially diminish this problem as data would be used as long as individuals do not explicitly indicate that they do not want their data to be used. A third option is to rely on data access regimes that include oversight bodies (Ausloos et al., 2020), such as *Findata* in Finland (Ausloos et al., 2020).

However, we know little about which of these transmission principles are more accepted across public benefit contexts (for context-specific research see, for example, on using phone data during the COVID-19 pandemic: Office of the Australian Information Commission & Lonergan Research, 2020). Given the scarcity of cross-context research, I formulate an open research question on this parameter:

RQ1: Which modes of consent do individuals accept more than other modes for data use for public benefit?

International Variation

As argued above, evaluations of social contexts, and thereby the effectiveness and acceptance of international policies surrounding these contexts, may vary by country. Previous privacy research has paid particular attention to cultural differences between countries by drawing on Hofstede's cultural dimensions (see Hofstede et al., 2010), which may also shape acceptance of data use for public benefit. However, one should be cautious to assume that there was a "national culture" permeating all domains of social life consistently (Masur et al., 2024). Different social contexts may have their very own privacy-related norms (Nissenbaum, 2010) that are not fully determined by general cultural orientations.

Among the cultural dimensions that pertain to the Hofstede approach, some scholars assess the individualism dimension to be the most central dimension for privacy by which to compare cultures (as discussed in Liu, 2022). Empirical research frequently identified effects of individualism on privacy-related phenomena. For instance, an international survey experiment found that public benefit uses of data are relatively more accepted than other uses by individuals with a more collectivist cultural background as compared with individuals with a more individualist cultural background (Li et al., 2017). The study also found that individualism is related to stronger effects of the option of "notice and control" methods on acceptance, and to lower acceptance of government as data collector. In a similar vein, another study on contact tracing apps found higher use willingness in China (which is considered rather collectivist) than in Germany and the United States (which are comparatively more individualist; Utz et al., 2021). However, other

studies provided an opposite relationship or null findings for individualism (see Engström et al., 2023; Trepte & Masur, 2016). Beyond individualism, less consistent or null effects have been found for the dimension of uncertainty avoidance (Engström et al., 2023; Schumacher et al., 2023; Trepte et al., 2017).

Given these findings, higher levels of individualism in a country may be associated with a higher desire of transmission principles that allow the affected individual more control over data flows. However, international differences may be hard to pinpoint to individualism with few countries of comparison, as countries may differ in further respects. Under this circumstance and partly contradictory findings, I approach international differences with an exploratory research question:

RQ2: Do countries with higher levels of individualism desire higher levels of control over data flows?

Interindividual Variation

Additional to the macro- and meso-levels, there may also be variation on the micro-level in assessing data use for public benefit, that is, individual differences within and between countries. Attitudes on data use for public benefit can relate to either of its constitutive elements of *data use* (i.e., privacy attitudes) and *public benefit*. Concretely, I distinguish between four types of relevant individual-level variables: (1) general attitudes and perceptions related to privacy, (2) perceptions with respect to specific elements of data flows, (3) general attitudes and perceptions related to the provision of public benefits, and (4) affinity toward technology and socio-demographic variables. First, individuals may differ with respect to privacy concerns, for instance, due to personality characteristics and own privacy-related experiences (Smith et al., 2011), and with respect to how they value privacy. The acceptance of data use scenarios may vary with individual general privacy concerns—possibly mediated by scenario-specific perceptions—(Kehr et al., 2015) and, second, with general perceptions relating to the parameters of the specific scenario: trust in data recipients (Kao & Sapp, 2022; Trein & Varone, 2023) and perceived sensitivity of data types (Mothersbaugh et al., 2012). Third, data use specifically for public benefit may be more accepted among individuals who value such public benefits higher more generally (relatedly for issue importance: Trein & Varone, 2023) and who have a more positive relationship to, or picture of, society, for example, who are more altruistic (Y. Kim & Stanton, 2016; Silber et al., 2022) and have higher interpersonal trust. Fourth, given the focus on digital data collection in the present article, familiarity with digital technologies may also affect privacy perceptions (e.g., Park, 2013). Finally, age and gender may affect privacy perceptions (Schomakers et al., 2019):

RQ3: Does the overall level of acceptance of data use scenarios vary with age, gender, general privacy concerns, perceptions relating to specific flow parameters, perceptions on

public benefit uses of data, altruism, interpersonal trust, and with affinity toward technology?

Longitudinal Variation

Finally, comparisons on all three levels (macro, meso, and micro) are contingent on the specific time point of the comparison. Perceptions on the importance and salience of privacy may change with major societal events (Büchi et al., 2022). Given that privacy has been intensely discussed during the COVID-19 pandemic, for example, due to contact tracing, privacy perceptions related to public health may have changed. A previous study found that acceptance of use of health data from smartphones for disease containment has increased from 2019 to spring 2020, but not for other non-pandemic-related data use scenarios, which supports the notion of context-dependent effects of societal developments (Gerdon et al., 2021). This notion has also found support with a longitudinal study on privacy attitudes in the United States, which has shown that acceptance to use fitness tracker data for medical research increased from 2019 to 2020 and then stayed higher, while this was not the case with government collecting data to counter terrorism, toward which acceptance decreased (Goetzen et al., 2022). However, salience or its effect may wane over time: Wnuk et al. (2021) conducted a longitudinal study in Poland and found that acceptance for (partly rather intrusive) surveillance technologies decreased between May and December 2020, that is, during the COVID-19 pandemic. This tendency did not change even when the pandemic threat was particularly high in the second wave of the pandemic.

Research therefore suggests that privacy attitudes related to health data may vary with the severity of the pandemic situation. However, it is unclear how severe the shifts in societal circumstances need to be to affect privacy attitudes. I compare developments in attitudes toward health data use with attitudes toward the use of other data types that are either also affected by current public issues (energy use data) or that are less immediately affected by current public issues (location and social media data):

RQ4: Does the acceptance of health data use, relative to the acceptance of using other data types, change with the pandemic situation?

Method

Experimental Design and Questionnaire

To compare a variety of data use scenarios within multiple social contexts, an online survey experiment (“vignette experiment,” see Auspurg & Hinz, 2015) was conducted. In this experiment, people’s attitudes toward several hypothetical scenarios in which data are used for the purposes of research and public management were measured. This experiment entails 33 text descriptions (so-called “vignettes”) of hypothetical data use

Table 2. Experimental Design: Vignette Factors and Levels.

Factors	Levels	Text
Data type ^a	<i>Health</i>	“a person’s health, diseases, and treatments”
	<i>Location</i>	“location of smartphones”
	<i>Energy use</i>	“the energy consumption of household appliances”
	<i>Social media</i>	“a person’s social media usage (for example Facebook and Twitter)”
Data recipient ^b	<i>University researchers</i>	“university researchers”
	<i>Researchers at an Internet company</i>	“researchers at an Internet company”
	<i>Public agency</i>	Further specified according to data type, e.g., “local public planning agency”
Transmission principle ^c	<i>Opt-in</i>	“... [recipient] may use this information for this purpose only if [data subject] agrees ...”
	<i>Opt-out</i>	“... [recipient] must not use this information for this purpose in any case if [data subject] rejects ...”
	<i>Ethics board</i>	“This information is stored at a national data storage centre. The [recipient] need[s] to request this information from this centre. A committee of independent ethics experts decides on the request.” + opt-out text

^aThe data subject is described depending on the data type (e.g., “resident” for energy use data).

^bCompany recipients were always defined as “researchers at an Internet company,” as these can be associated with handling different types of data for various purposes. As there is no public agency recipient that could be directly associated with such a multitude of data uses, the vignettes refer to different specific public agencies depending on data type. The two public and the one private recipient come with different data use purposes to create realistic scenarios. The purpose for university and company researchers always is research, and for agencies it is planning or control. For instance, researchers (both public and private) may use health data to study diseases. Public agencies are presented to use data for planning or control purposes, e.g., location data for infrastructure planning.

^cTo describe the “ethics board,” I refer to “committee of independent ethics experts” and “data centres” as simplifications that work across data types. The exact means to either accept or reject the data use is adjusted to data type to increase plausibility. The basis for all descriptions of the transmission principles is that individuals are informed about the data use.

scenarios and allows researchers to estimate how changes in scenario characteristics affect acceptance. The vignettes vary by factors that can take on different levels (Auspurg & Hinz, 2015), with the factors representing contextual integrity parameters.³ Table 2 shows the factors and levels. The full list of vignettes is available in the Supplementary Materials (*Section V*).

Each respondent was presented with four vignettes, receiving exactly one random vignette for each of the four data types. The order of shown data types was random, with one exception: As the social media vignettes did not contain all possible combinations of vignette factors (see Note 3), these vignettes were always placed in the last position and I treat them as a separate experiment in order to maintain a fully factorial experimental design (see Auspurg & Hinz, 2015) for the other vignettes. Respondents were asked to evaluate each presented scenario by rating the appropriateness of the data use on a seven-point scale.⁴ Afterwards, respondents were shown items that measure relevant concepts for the research questions on interindividual variation.⁵ The questionnaires are available in *Section Q* and information on cognitive pre-tests and a pilot study is available in *Section M* of the Supplementary Materials.

Sample

The vignette experiment was conducted as part of an online survey. The survey was fielded in three countries that varied in their levels of individualism according to the Hofstede

cultural values index (Hofstede Insights, 2023): Germany, Spain, and the UK. According to Hofstede’s dimensions, the UK is a state with a high level of individualism, while Germany displays medium values (Hofstede Insights, 2023). Spain is one of the countries with the lowest individualism scores in Europe (Hofstede Insights, 2023).

Respondents were invited to participate via a commercial non-probability online panel provider (*Bilendi*). Respondents can self-select into the respondent pool and are then invited and incentivized by the provider to participate in specific surveys. For this survey, crossed age and gender quotas were applied that correspond to the respective population distributions based on Eurostat data from 2020. While inference with non-probability samples is oftentimes problematic (Elliott & Valliant, 2017), the goal is to estimate effects of experimental stimuli, which non-probability samples can be useful for (Jamieson et al., 2023; Kohler & Post, 2023), and to explore associations with individual-level variables.

The survey was fielded at two time points: 14 to 21 December, 2022 (*Wave 1*) and 11 to 22 May 2023 (*Wave 2*).⁶ Based on these two cross-sectional samples, I also constructed a longitudinal data set of respondents who participated in both waves. In the second wave, each recurring respondent received the same vignettes in the same order as in the first wave.

The sample sizes for analysis are as follows: Wave 1: 1,682 respondents (Germany: 562; Spain: 564; UK: 556); Wave 2: 1,795 respondents (Germany: 594; Spain: 603; UK: 598). Wave 2 comprises 1,110 respondents who already participated

Table 3. Overview of Used Regression Models.

Model	Predictors
Model 1	vignette levels, vignette positions, country dummies, age, gender; random intercept for the respondents
Model 2	Model 1 + interaction between recipient and data type
Model 3	Model 1 + interaction between recipient and transmission principle
Model 4	Model 1 + individual-level variables
Model types	<i>Main model:</i> Linear mixed-effect models (<i>lmer</i>) without speeders <i>Alternatives:</i> Linear mixed-effect models (<i>lmer</i>) with speeders “Ordinal” logistic cumulative link mixed models (<i>clmm</i>) For social media data: OLS models (<i>lm</i>) and ordinal models (<i>clmm2</i>) without random intercepts

I analyze responses to the “social media” vignettes as a separate experiment (see “Method” section) using OLS regression. I ran two additional types of models to check whether the findings are robust to model and data choices: (1) models that include respondents that I defined as speeders and (2) logistic cumulative link mixed models (Christensen, 2019) that treat the outcome variables as ordinal. I focus on the interpretation of the linear mixed-effects models and highlight important substantive differences compared with the two other types of models. Note that I cannot directly compare effect sizes of the ordinal models with other models (Mood, 2010).

in wave 1; the remaining number of participants was newly recruited.⁷ Details on the sample and exclusion criteria are available in the Supplementary Materials (*Section M*).

Results

To address the hypotheses and research questions on contextual and international variation, I use the data from Wave 1 and regress perceived appropriateness on the vignette variables and on country dummy variables, while controlling for age and gender. To this end, I run linear mixed-effects models with random intercepts for the respondent-level. I investigate further context-dependencies of effects of data recipients by adding respective interaction terms. To answer the research question related to interindividual variation, I add individual-level variables to the regression models. For the longitudinal comparisons, I inspect changes in acceptance (note that I will use the terms “acceptance” and “perceived appropriateness” interchangeably for easier text flow) of specific scenarios from Wave 1 to Wave 2. Finally, I use the Wave 2 data for a replication of the Wave 1 results by running the same regression analyses as in Wave 1 and inspecting whether substantive changes occur. Table 3 summarizes the models that will be shown in the “Results” section.⁸

Note that although I draw on a non-probability sample, I calculate standard errors as an orientation using the, respectively,

implemented procedures of the software. However, the focus is on the interpretation of effect strength.

Age and gender distributions (Supplementary Table T1) and summary statistics for all other individual-level variables are available in the Supplementary Materials (Table T2a for Wave 1 and Table T2b for Wave 2).

Contextual and International Variation in December 2022

Before turning to the regression analyses, I descriptively inspect the mean values of the vignette scenarios across countries in Wave 1. The mean perceived appropriateness varies across countries and vignette factors (Figure 1). While the ratings do not vary strongly for some vignettes and between single vignette levels, some patterns are discernable. The highest perceived appropriateness is found for health and energy data, followed by location data and then by social media data. However, the lower acceptance of social media vignettes could be partly driven by always being the last vignette to be shown to respondents (see below). Finally, while not true for each single scenario, respondents from the UK appeared as the overall most accepting country, followed by Spain and then Germany. The Supplementary Materials contain the exact mean and median values (Table T3a) and the full distribution of answers for each vignette (Figure F1a).

To answer H1, H2, and RQ1 (the effects of vignette factors on acceptance), I compare the effects of vignette factors on perceived appropriateness ratings across countries. To this end, I run regression analysis with pooled data from all countries as well as separately for each country.

I first run linear mixed-effects models that only contain the vignette levels, vignette positions, country dummies, age, gender, and a random intercept for the respondents (Figure 2; M1 columns in Table T4a in the Supplementary Materials; ordinal and speeder models in Tables T4b and T4c). The results show that among vignette factors, the data types have the strongest effects. Scenarios with health data overall appear as more accepted than energy use data, while location data use is rated lower. However, in the UK, there are no meaningful differences between health and energy use data (albeit the effect is slightly stronger in the ordinal model). These findings support H1 (“health data use is more accepted than the use of other data types”) in Germany and Spain, while for UK, the use of health and energy data is similarly accepted.

While respondents overall do not strongly differentiate between company and university researchers, there is a slight preference for the latter in Germany and Spain. Moreover, except for in Spain, respondents rate vignettes with public agencies similar or lower than vignettes with researchers. This rejects H2 (“public actors, and particularly public researchers, are more accepted than private actors”) overall, there being only a slight such tendency in Spain. To learn

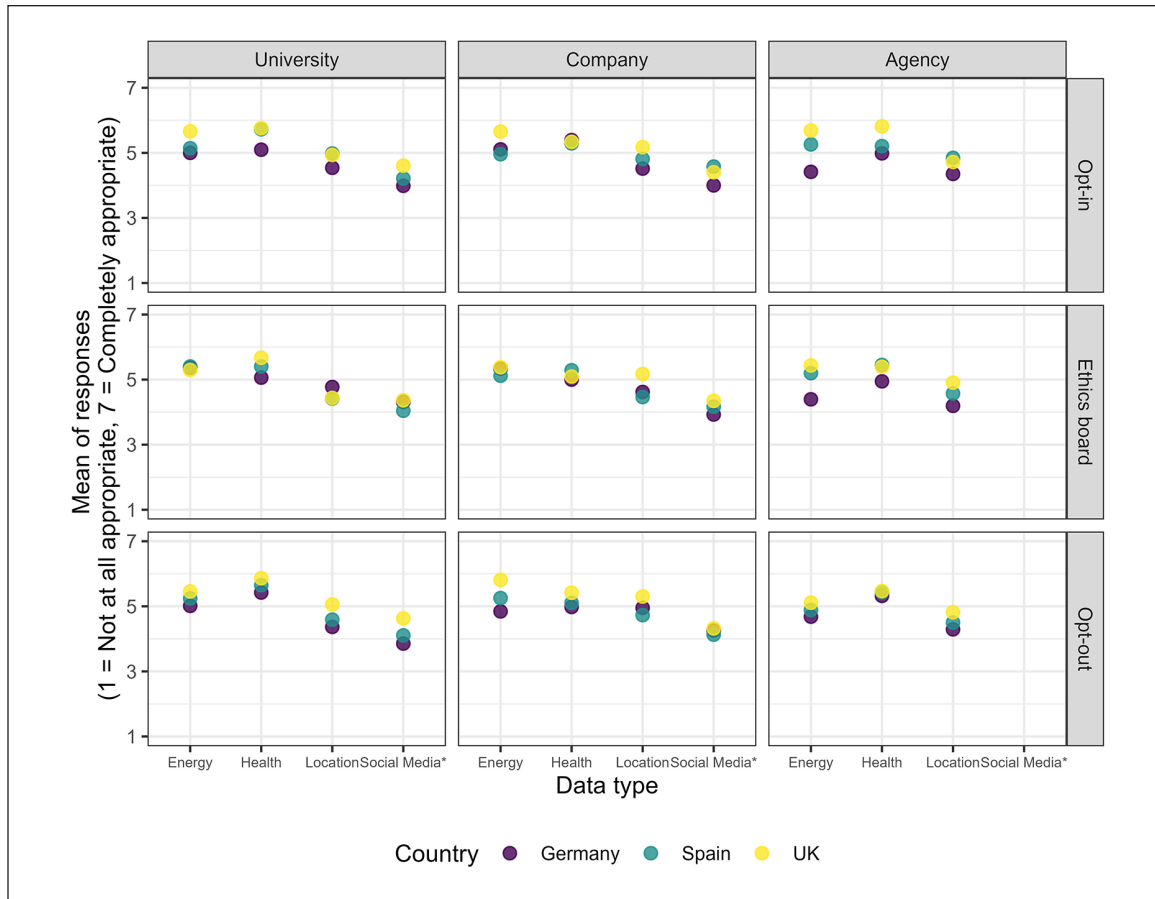


Figure 1. Arithmetic mean values of perceived appropriateness of all vignette scenarios in Wave I (December 2022).

Each column represents one data recipient, each row one transmission principle. Each box shows the arithmetic mean values for each data type and for each country. Number of responses per country: Germany: 2,248; Spain: 2,256; UK: 2,224.

*Note that the social media vignettes were always shown last and are treated as a separate experiment, such that their overall lower acceptance values might at least partly be due to order effects.

whether the recipient effect varies by data type and transmission principles, I add the respective interactions in two separate models. That is, there is one model with interaction between recipient and data type (Model 2) as well as one model with interactions between recipient and transmission principle (Model 3). All results are available in Table S1 in the Supplementary Materials (columns M2 and M3).

The results for Model 2 show that many interactions between recipients and data types are likely random—given the oftentimes small effect sizes and large standard errors—but there are some stronger effects. There is a positive interaction effect between agencies and health data in Germany. Agency recipients are in tendency more accepted for location data than for energy use data. Companies are in tendency less accepted for health data and more accepted for location data, compared with energy use data; however, for companies, there are barely such differences in Germany.

The results for Model 3 show mostly small and likely random interaction effects between recipients and transmission principles. Two of the more consistent findings are that the

combination of an agency recipient and opt-out compared with the reference categories is somewhat less accepted in the UK, and that the combination of company and ethics board is somewhat more accepted in Spain, again compared with the reference categories.

In summary, the results confirm the expectation that recipient and data type interact. There is less consistent evidence for strong interactions between recipient and transmission principle, although somewhat stronger effects show for specific combinations.

Finally, I analyze the additional experiment on social media data (Table T5a in the Supplementary Materials; ordinal and speeder models in Supplementary Tables T5b and T5c). The social media vignette was always placed in the last position and never contained a public agency as a recipient. The results show that most effects could be random, but there are some tendencies. Again, overall acceptance is higher in UK and Spain than Germany, but the latter difference is smaller than in the previous models. Depending on the country, respondents assess company researchers differently, as

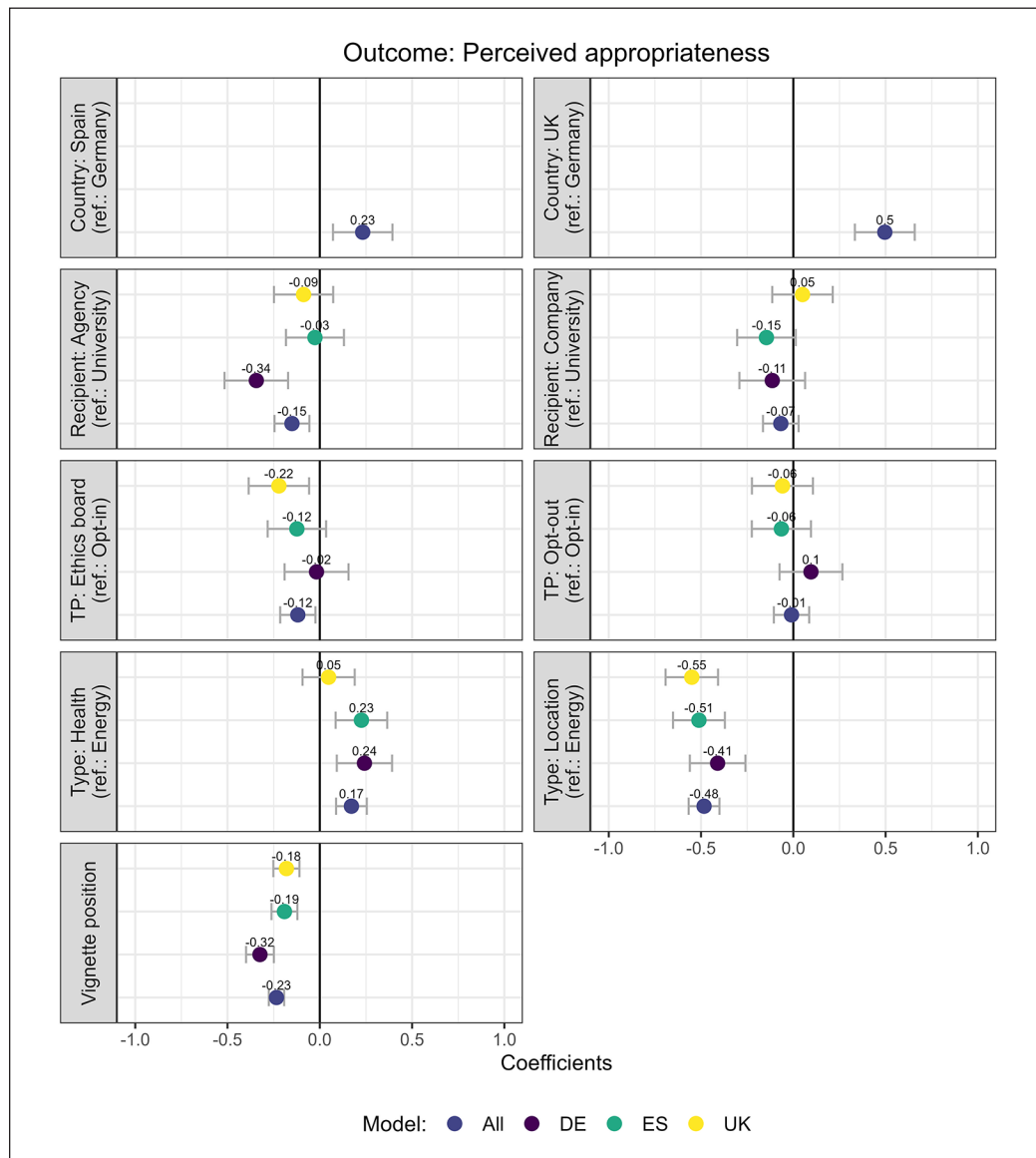


Figure 2. Linear mixed-effects model regression coefficients and 95% confidence intervals for effects of vignette levels on perceived appropriateness in Wave I, based on four separate models.

TP means “transmission principle.” N: All: 4,562; Germany: 1,510; Spain: 1,549; UK: 1,503. Models further include age, gender, and a random intercept on the respondent-level (not displayed in the figure).

compared with university researchers. In Germany, there are no strong differences (except for a somewhat stronger negative effect in the model with speeders), while there is a higher relative acceptance in Spain and a lower relative acceptance in the UK. The effects of transmission principles in tendency vary by data recipient and across countries (see Supplementary Table T5a for details).

Based on these results, I can answer RQ1 (“Which modes of consent are more accepted?”). The above models show no overall strong differences between opt-in and opt-out procedures (except for a somewhat stronger negative effect for opt-out in the case of social media data in Spain). Ethics boards are in tendency less accepted than opt-in procedures. Particularly the latter effect varies by country.

This leads to RQ2 (“Do countries with higher levels of individualism desire higher levels of control over data flows?”). While acceptance in all three countries barely changes between opt-in and opt-out procedures, UK respondents are somewhat relatively more skeptical about ethics boards. Spanish respondents accept ethics boards slightly less than opt-in. If individualism was responsible for international differences, there should be clearer differences in the acceptance of transmission principles (especially for opt-out vs opt-in) particularly between UK and Spain. Moreover, contrary to expectation, the overall acceptance is highest in the UK. The answer to RQ2 thus is that there is no clear and consistent association of a higher desire for control with higher country-level individualism.

Interindividual and International Variation in December 2022

To answer RQ3 on the associations of individual-level variables with perceived appropriateness, I add variables to Model 1 that are related to trust, altruism, perceived sensitivity of data types, other privacy-related perceptions, device use, and affinity toward technology.

I focus on the effects of individual-level variables across models (Figure 3; Table T6a in the Supplementary Materials; ordinal and speeder models in Supplementary Tables T6b and T6c). On average, female respondents report somewhat lower acceptance than male respondents. Four associations are relatively consistent across countries: higher trust in data recipients comes with higher acceptance, while higher perceived sensitivity of data types comes with lower acceptance (note that sensitivity and trust always refer to the specific data type or data recipient shown in the vignette). General privacy concerns are associated with lower acceptance, while agreement with the statement that “The privacy of individuals may be invaded if this results in a greater benefit to society” (own translation from Trepte, 2020) comes with higher acceptance. Additional to these consistent associations, more thinking about privacy in times of the pandemic is associated with higher acceptance in the UK. Altruism has a positive association with acceptance, particularly in Germany. The same is true for the number of used devices in Spain. Otherwise, there are mostly small and likely random associations (although single effects appear more meaningful in the alternative model types). With respect to the effects of vignette factors, it is noteworthy that controlling for the individual-level variables, public agencies appear as the most accepted recipient in Spain, while company recipients are most accepted in Germany and the UK.

Longitudinal Variation

To answer RQ4 (“Does the acceptance of health data use, relative to the acceptance of using other data types, change with the pandemic situation?”), I first inspect changes in perceived appropriateness from Wave 1 (December 2022) to Wave 2 (May 2023) across data types and countries. To this end, I use a data set that only contains respondents who participated in both waves (see details on the construction of this sample in Supplementary Materials *Section M*). The Supplementary Materials contain mean values—plotted (Figure F2) and as a table (Table T3b)—and the full distribution (Figure F1b) of vignette responses in Wave 2.

When analyzing longitudinal changes in acceptance separately for each data flow parameter or by country, only small differences over time are observed (Figure 4; with speeders: Figure F3 in Supplementary Materials). A comparison of changes between more specific vignette scenarios reveals a more nuanced picture (Figure 5; exact values in Table T7a in Supplementary Materials). Acceptance changed only slightly

for several combinations of data types and recipients. However, in the UK, health data use became less accepted for companies and more accepted for universities. Changes for energy and social media data are almost consistently negative and, in some cases, relatively small. Moreover, there are some stronger increases in acceptance for location data use. With respect to RQ4, these results still show that the acceptance of health data did not overall decrease (or increase) much more relative to other data types. While there are stronger changes for health data with specific recipients in the UK (and partly Spain), there are similarly strong changes for other settings (when including speeders, however, the decrease for company recipients in the UK is particularly strong, but the increase for public agency is less pronounced; see Figure F4 and Table T7b in the Supplementary Materials). However, some changes vary considerably when further taking into account transmission principles (but note that the number of responses per combination is lower in this more fine-grained analysis). In some cases, the ratings of the same data type and recipient changes into different directions depending on transmission principles (Supplementary Table T8a and Supplementary Figure F5a without speeders, Supplementary Table T8b and Supplementary Figure F5b with speeders). Still, ratings of health data vignettes do not stand out to have overall changed particularly more than other ratings.

Replication of December 2022 Results With Data From May 2023

Finally, I make use of the full data set for Wave 2 that comprises respondents who already participated in Wave 1 as well as newly recruited respondents. I treat this second wave as a replication of the first wave and check whether the substantive findings with respect to the hypotheses and research questions hold. However, this approach cannot reveal whether any differences are attributable to changes over time or to differences in sample composition.

I run all models from Wave 1 again with data from the second wave and show all regression tables in the Supplementary Materials (Tables with ending letters *d* to *f*). The finding holds that health data are more accepted than the other data types (H1). In fact, in Wave 2, there are somewhat stronger positive effects for health data for the UK, compared with Wave 1. Also, public recipients (H2) are again not overall clearly more accepted than company recipients, although the relative acceptance of the latter in tendency is lower. As for transmission principles and their importance across countries (RQ1 and RQ2), opt-out is the overall most accepted transmission principle in Spain in Wave 2 (except when using social media data), while opt-in and opt-out are again most accepted in the other countries. The individual-level variables (RQ3) overall display the same tendencies as in Wave 1, but women display rather equal acceptance compared with men in the UK. However, the associations with interpersonal trust and of

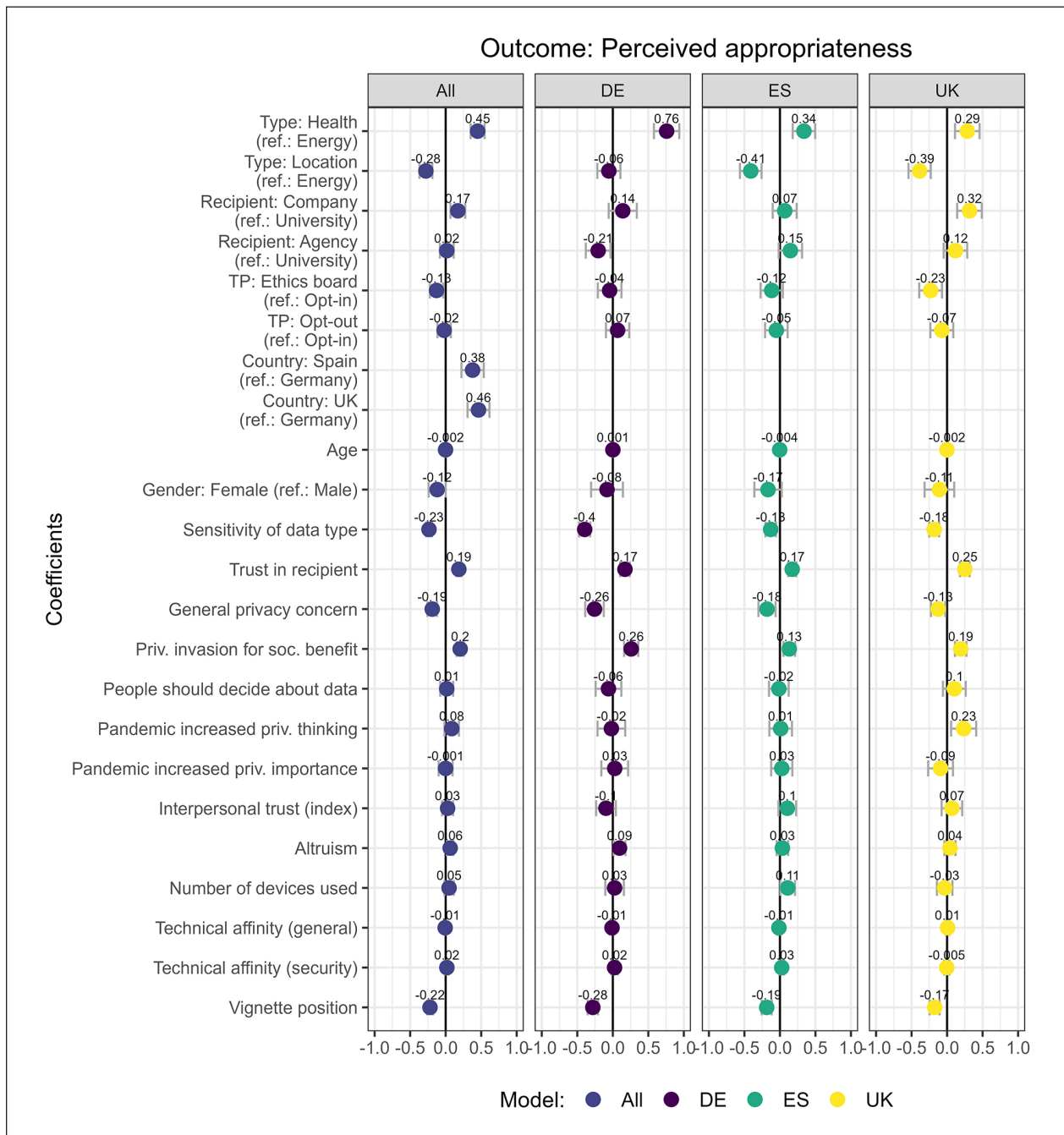


Figure 3. Linear mixed-effects model regression coefficients and 95% confidence intervals for effects of vignette levels and individual-level characteristics on perceived appropriateness in Wave I, based on four separate models. TP means “transmission principle.” N: All: 4,562; Germany: 1,510; Spain: 1,549; UK: 1,503. Models further include a random intercept on the respondent-level (not displayed in the figure).

thinking about privacy in times of the pandemic tend more toward zero. Moreover, the differences between recipients tend to be somewhat smaller, a slight exception being a relatively higher acceptance of agency recipients in the UK than in Wave 1. The resulting cross-country patterns may be somewhat more in line with the differences in levels of individualism in the respective countries but are still not clearly consistent and pronounced. Across models, there are changes in effects

for further specific constellations—especially for the interaction effects and for the case of social media data—that can be ascertained from the respective tables.

Discussion

The results demonstrate considerable variation of perceived appropriateness of data use for public benefit across

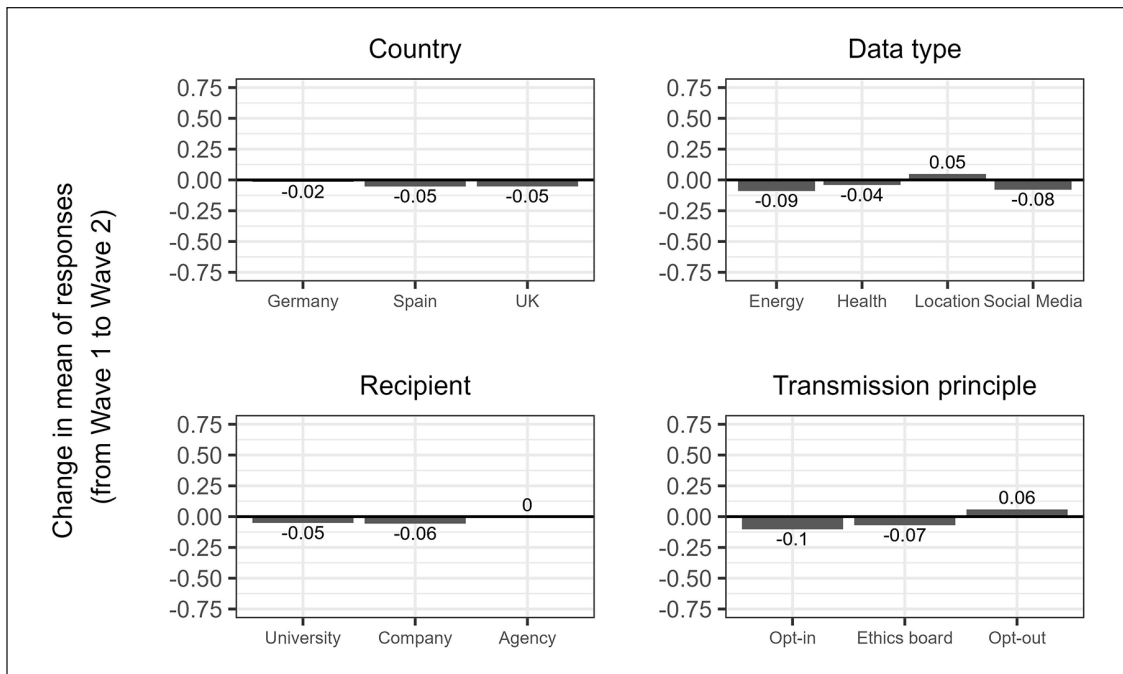


Figure 4. Changes in arithmetic means of responses from Wave 1 (December 2022) to Wave 2 (May 2023) among those respondents who participated in both waves. Aggregated for country, data type, recipient, or transmission principle. Based on 8,880 responses from 1,110 respondents who participated in both waves.

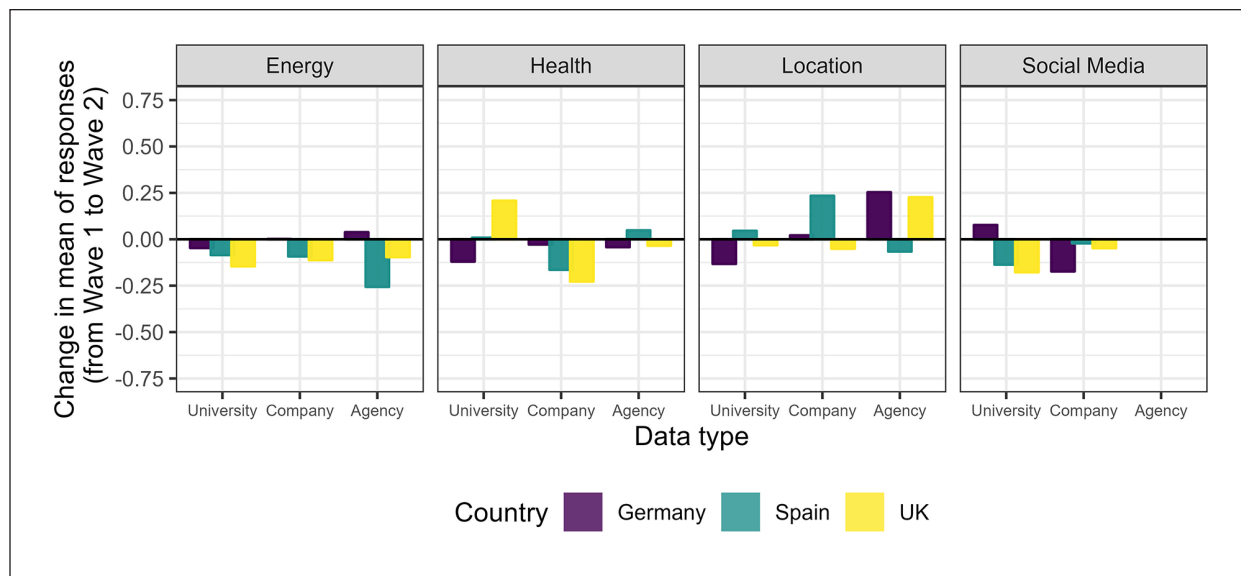


Figure 5. Changes in arithmetic means of responses from Wave 1 (December 2022) to Wave 2 (May 2023) among those respondents who participated in both waves. Differentiated by country, data type, and data recipient. Based on 8,880 responses from 1,110 respondents who participated in both waves.

contexts. The contextual effects, moreover, vary by country and, to some extent, over time. These findings support the notion that contextual integrity is a useful approach for comparative research across countries: Additional to country comparisons with respect to *general* privacy notions, contextual integrity can reveal *context-specific* differences

between countries that may otherwise remain unnoticed (while this gap between general and specific perceptions has already been argued for within single countries, Martin & Nissenbaum, 2017). This study integrates contextual integrity into comparative privacy research (see Masur et al., 2024) and demonstrates that future research can

operationalize data uses in meso-level social contexts by drawing on contextual integrity's data flow parameters. The results also show that some individual-level variables are mostly consistently associated with higher or lower acceptance. In the following, I discuss more specific implications and research avenues with respect to contextual, international, interindividual, and longitudinal comparisons, before turning to limitations of the study.

Among contextual factors, changes in data types had the strongest effects on acceptance. The relatively high acceptance of health data use is somewhat striking as one might consider this data type to be particularly sensitive. Indeed, a higher sensitivity of data types is associated with lower acceptance, but the positive estimates for health data remain. This finding supports the notion that sensitivity is a context-dependent concept (Martin & Nissenbaum, 2017). The relatively higher acceptance of health and energy use data, compared with location data, could be explained by their particular societal relevance, while the lower acceptance for social media vignettes may be due to an order effect. Further longitudinal research would need to investigate whether this represents a temporary or stable difference in preferences.

Importantly, the findings highlight the relevance of contextual integrity parameters for international comparisons. For instance, public agencies are relatively less accepted data recipients in Germany compared with the other countries. These findings have two implications for future comparative privacy research. First, country comparisons of meso-level social contexts can identify cross-country differences that are not captured by general differences in privacy perceptions, for example, in health versus energy use contexts. Second, data use contexts that appear similar across countries may be differently evaluated by the respective populations, calling for further cross-country research using the notion of contextual integrity. For instance, while acceptance toward data use by public agencies may appear overall relatively lower in Germany, the acceptance of this data recipient still varies by used data type within Germany: As data types were particularly influential, separate ordinary least squares (OLS) models for each country and data type in the Supplementary Materials further illustrate data type- and country-specific differences (Tables T9a to T9f) for interested readers.

The found differences across countries are not clearly and consistently explainable by country-level individualism, although the results from Wave 2 resemble the expected patterns somewhat more than those from Wave 1. Future research may further context-specifically investigate cultural dimensions that better explain the found differences. For instance, Germany scores high and the UK scores low on Hofstede's dimension of "uncertainty avoidance," with Spain being ranked between these two countries (Hofstede Insights, 2023). However, it may be that country-level cultural variables cannot capture the complexities that are inherent to international comparisons. Scholars criticized

Hofstede's cultural dimensions approach for various reasons, among them methods-related concerns and considering it a too positivist approach (discussed in Jackson, 2020). As the Comparative Privacy Research Framework suggests, countries may differ with respect to a variety of categories, not only the cultural aspect of individualism (Masur et al., 2024). To learn more about the concrete categories that matter, future research would need to include a large variety of countries that differ with respect to multiple categories at different levels, such as the *economic* category (see Masur et al., 2024). For example, degrees of free market economy (Masur et al., 2024) could be related to the establishment of relatively free use of data in the respective countries. However, while country-level characteristics may explain some of the differences between countries, the present study suggests that a context-specific view is necessary to avoid undue generalizations to all kinds of data uses.

The analysis further explored associations between individual-level variables and acceptance. The results show that even after taking into account contextual factors, several general privacy-related perceptions are still associated with acceptance. Adding to Martin and Nissenbaum's (2017) suggestion that context-specific preference measurements may partly close the gap between stated privacy concerns and situation-specific data sharing behavior (the so-called "privacy paradox"), these findings imply that the measurement of general perceptions may still be worthwhile in context-based research. Future research needs to investigate how context-specific alongside general privacy perceptions translate into data sharing behavior.

Some changes in perceived appropriateness over time were found in the longitudinal comparisons of specific contextual constellations, but not pronouncedly and universally for the acceptance of health data use. It might be that respondents were on average not overly concerned about the COVID-19 pandemic anymore already at the time of data collection in December 2022. As for energy supply problems as another salient public issue, there is a tendency toward decreased acceptance from December to May, but these changes are overall not very strong. However, given some stronger changes for specific constellations, these results highlight again that privacy research and policymaking need to reflect how the timing of data collection—for example, during a specific crisis—might affect context-specific results. Momentary assessments of public opinion do not necessarily constitute a "social license" (Shaw et al., 2020) to carry out questionable data uses. Instead, they constitute one element of an assessment of the appropriateness of a data use, along with further legal considerations and, as suggested by contextual integrity, the discussion of context-specific and more general values and goals at stake (Nissenbaum, 2010).

I now turn to limitations of the article that were not already discussed above. First, as is commonly the case with vignette studies, we need to keep in mind potential limitations with

respect to external validity (Eifler & Petzold, 2019). Moreover, this study can only speak about the concrete investigated scenarios. For instance, the company recipients were always defined as “researchers at an Internet company” and it is possible that recipients rate vignettes differently if more context-specific companies are involved. The relatively strong effects found for data types may also be due to the circumstance that the vignettes were structured around data types to make them appear plausible and to not present, for example, construction planning agencies using patient records. Moreover, some of the investigated scenarios may have been very hypothetical or unknown to respondents. With increasing concreteness or public awareness, attitudes toward these data uses may change.

Second, in principle, internationally different response behavior patterns (Kemmelmeier, 2016)—however, less so due to the experimental design—and potential variations in the interpretations of vignettes may account for some of the differences found between countries. Researchers also need to validate the cross-cultural invariance of privacy-related measurements (Ghaiumy Anaraky et al., 2021). Moreover, a larger number of countries might allow researchers to better disentangle effects of, for example, cultural and economic differences (see Masur et al., 2024) on privacy attitudes. While this study has detected differences even between three European countries, differences could be further pronounced particularly when extending comparisons to non-WEIRD—that is, western, educated, industrialized, rich, and democratic (Henrich et al., 2010)—countries.

Third, as explained above, the study is based on an online non-probability sample for which inference is only feasible under specific conditions and for specific fields of application (Kohler & Post, 2023). This means that while more confident claims with respect to experimental effects can be possible, for example, mere mean values are not to be inferred to the general populations of the respective countries. Moreover, as the questionnaire was fielded online, one could speculate that particularly older respondents in this sample might be more open to digital technologies than older individuals in the general population, which could explain the small or non-associations of age with acceptance. Thus, future research would need to confirm these findings with mixed-mode or offline probability samples.

Conclusion

The Comparative Privacy Research Framework (Masur et al., 2024) proposes to compare privacy-related phenomena across different levels and types of units. The present study draws on the concept of “contextual integrity” (Nissenbaum, 2010) to enhance comparative privacy research by focusing on social contexts as a meso-level unit of comparison. To this end, the contextual integrity data flow parameters offer a useful approach to operationalize data uses in meso-level social contexts. The present study applies this approach by employing a

survey experiment in three countries (Germany, Spain, and the UK) and at two time points (December 2022 and May 2023) to compare privacy perceptions related to data use for public benefit along four components: contextual, international, interindividual, and longitudinal variation. The results show that the effects of data flow parameters vary across countries, but to different degrees. The strongest effects are found for the data type, with health data overall being the relatively most accepted data type overall. The effect for data recipients varies such as to lead to different substantive conclusions for the different countries. Country-level individualism was not found to be clearly and consistently associated with the desire for control over the data. Interindividually, several general privacy perceptions still matter after considering contextual factors. Finally, longitudinal comparisons show overall minor but context-dependent variation over time.

In conclusion, using the contextual integrity approach can unmask meso-level context-specific differences in the acceptance of data uses within and between countries. These differences could be relevant for ascertaining “social licenses” (see Shaw et al., 2020) regarding data use practices, for international privacy-related regulation, and for suggestions on sector-specific policies. This study can therefore serve as a call for more deliberately incorporating meso-level contexts in comparative privacy research that can inform privacy-related public decision-making.

Acknowledgements

The author thanks Ellen Laurischk, Kerstin Fischer, and Tonja Dingerdissen for their support in this research project, and Elsa Peris, Wiebke Weber, and Joshua Fullard for their work or support regarding questionnaire translation. He also thanks Frauke Kreuter, Helen Nissenbaum, and Thomas Fetzer for their comments and ideas, and the Kreuter-Keusch research group and the reviewers for helpful comments on this manuscript. A previous version of this article is part of the publication-based dissertation of the author (URN: urn:nbn:de:bsz:180-madoc-674360).

Data Availability

It is planned to publish the data in a safe data repository for scientific use. Until then, the data are available from the author upon request.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by Volkswagen Foundation (Grant “The Covid-19 Pandemic and Data Sharing for the Public Good: Attitudinal, Ethical, and Legal Approaches to Privacy During the Pandemic and Beyond”). Further funding for data collection came from the Ludwig-Maximilians-Universität München. This work was supported by the University of Mannheim’s Graduate School of Economic and Social Sciences.

Ethical Approval and Consent

In Germany, research that does not involve patients usually does not require approval by an ethics committee. Yet, approval for this research endeavor was voluntarily applied for and obtained from the Ethics Committee of the Faculty of Mathematics, Computer Science, and Statistics at Ludwig-Maximilians-Universität München (approval: EK-MIS-2021-075, date: 4 February 2022; note that the surveys were initially intended to be conducted in Denmark and the United States instead of Spain and the UK). Respondents consented to participation on the first page of the questionnaire that contained information on the study and the use of data.

ORCID iD

Frederic Gerdon  <https://orcid.org/0000-0003-4442-6698>

Supplemental Material

Supplemental material for this article is available online.

Notes

1. Another example are mobility data (e.g., from smartphones) that can be used by companies for the personal benefit of drivers by suggesting optimal routes to destinations. These mobility data could also be used by researchers or public agencies to learn about mobility behavior on a fine-grained level for infrastructure planning that benefits the local population.
2. For instance, international differences in the acceptance of digital patient records could in principle deviate from international differences found for general privacy concerns or general perceptions toward health data, and the influence of who exactly would receive these data could vary across countries.
3. The vignettes are constructed around four data types: health data (digital patient records), location data (smartphone location), household energy use data (see Horne et al., 2015), and social media data. Particularly, health data have been very salient during the COVID-19 pandemic. To some extent, this may be true for location data as well as contact tracing apps sparked discussions on how to store information on contacts of individuals. Energy usage is unrelated to the pandemic but may have been particularly salient in winter 2022 due to public discussions about energy supply. Finally, social media data are of particular interest to researchers, but not directly related to the pandemic. As for data recipients, individuals may vary by their concerns about public and private data recipients, which is why public agencies get compared with researchers from universities and companies. Note that for the social media data type, I exclude public agencies as a data recipient as this might appear as a too intrusive scenario to respondents. Finally, I differentiate between three transmission principles under which individual data may be shared with the recipients: opt-in, opt-out, and a combination of opt-out with an ethics board (see section “Contextual variation”). Combining four data types, three recipients, and three transmission principles, and excluding governmental agencies as a recipient for social media data use, results in a total of 33 vignettes. For example, the vignette with the combination health data (data type), university researchers (recipient), and opt-in (transmission principle) reads thus:

Information about a person’s health, diseases, and treatments can be stored in a digital patient record.

Each person is informed in a doctor’s surgery about the possibility that university researchers could use this information to study diseases. This information does not contain the person’s name or address.

The researchers may use this information for this purpose only if the person agrees to this use verbally or in writing via a form after being informed.

4. To capture approval in general terms for scenarios that may not yet exist, or which respondents may not be aware of, the question was: “To what extent would you say that the use of the information as described above is appropriate or not appropriate?,” with a fully labeled seven-point scale ranging from *Completely appropriate* to *Not at all appropriate*.
5. This includes, in the following order: general privacy concerns (translated and slightly edited version from Trepte, 2020); trust in all possible vignette data recipients (based on ESS Round 9: European Social Survey, 2021); perceived sensitivity of all possible vignette data types (based on Pew Research Center, 2014); two items on agreement with statements relating to control over and public use of individual data (translated from Trepte, 2020); whether respondents think that their concerns about and importance of privacy changed from before to during the COVID-19 pandemic (loosely based on Office of the Australian Information Commission & Lonergan Research, 2020); interpersonal trust measured by an index based on three variables as presented in the *Interpersonal Trust Short Scale* (KUSIV3), where the sum is divided by the number of items (Nießen et al., 2021); altruism (from SOEP-IS Group, 2021); technical affinity, measured by the number of regularly used communication devices (based on Bauer et al., 2022); and additive indices of two subscales (*General and Safe application*) of the *Information and Communication Technology Self-Concept Scale* (ICT-SC25) to measure affinity toward technology (Schauffel et al., 2021).
6. The societal environment with respect to health and energy data perceptions potentially differs between these two time points, as December is a colder month where infectious diseases (such as COVID-19) and the energy crises overall may affect people’s lives more than in spring. The two time points are also not too far separated, which makes it less likely that other major events happen that systematically affect perceptions related to the vignette scenarios between the surveys.
7. To exclude potentially inattentive respondents from the analyses, these samples do not include “speeders,” that is, respondents who completed the questionnaire in less than 60% of the country- and wave-specific median response time (Roßmann, 2010). A small number of other respondents have been excluded for, for example, not agreeing to reading the questions carefully (see Conrad et al., 2017). The final distribution of age and gender for each country and wave is available in the Supplementary Materials (Table T1). The age distribution shifts toward a higher prevalence of older age groups in Wave 2 and the longitudinal sample as compared with Wave 1.

8. All models include only those respondents who have no missing or “Don’t know” (or similar) values for any of the individual-level variables that are introduced in the later extended models. However, I do not remove all vignette responses by one respondent from the data set if the respondent has missing or “Don’t know” values only for single vignettes. For data preparation, analysis, and presentation/visualization, I use *R* version 4.0.4 (R Core Team, 2021) with the libraries: *ggpubr* 0.4.0 (Kassambara, 2020), *grid* (R Core Team, 2021), *lme4* 1.1-29 (Bates et al., 2015), *ordinal* 2019.12-10 (Christensen, 2019), *sjPlot* 2.8.14 (Lüdtke, 2023), *tidyverse* 2.0.0 (Wickham et al., 2019), and *viridis* 0.6.0 (Garnier et al., 2021).

References

- Aitken, M., de St. Jorre, J., Pagliari, C., Jepson, R., & Cunningham-Burley, S. (2016). Public responses to the sharing and linkage of health data for research purposes: A systematic review and thematic synthesis of qualitative studies. *BMC Medical Ethics*, 17(1), Article 73. <https://doi.org/10.1186/s12910-016-0153-x>
- Ausloos, J., Leerssen, P., & ten Thijs, P. (2020). *Operationalizing research access in platform governance. What to learn from other industries?* https://algorithmwatch.org/en/wp-content/uploads/2020/06/GoverningPlatforms_IViR_study_June2020-AlgorithmWatch-2020-06-24.pdf
- Auspurg, K., & Hinz, T. (2015). *Factorial survey experiments*. Sage.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1), 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Bauer, P. C., Gerdon, F., Keusch, F., Kreuter, F., & Vannette, D. (2022). Did the GDPR increase trust in data collectors? Evidence from observational and experimental data. *Information, Communication & Society*, 25(14), 2101–2121. <https://doi.org/10.1080/1369118X.2021.1927138>
- Best, S. J., Krueger, B. S., & Ladewig, J. (2006). Privacy in the information age. *Public Opinion Quarterly*, 70(3), 375–401. <https://doi.org/10.1093/poq/nfl018>
- Büchi, M., Festic, N., & Latzer, M. (2022). The chilling effects of digital dataveillance: A theoretical model and an empirical research agenda. *Big Data & Society*, 9(1), 205395172110653. <https://doi.org/10.1177/20539517211065368>
- Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: Why care.data ran into trouble. *Journal of Medical Ethics*, 41(5), 404–409. <https://doi.org/10.1136/medethics-2014-102374>
- Christensen, R. H. B. (2019). *Ordinal—Regression models for ordinal data. R package version 2019.12-10* [Computer software]. <https://CRAN.R-project.org/package=ordinal>
- Conrad, F., Tourangeau, R., Couper, M., & Zhang, C. (2017). Reducing speeding in web surveys by providing immediate feedback. *Survey Research Methods*, 11(1), 45–61. <https://doi.org/10.18148/SRM/2017.V11I1.6304>
- De Man, Y., Wieland-Jorna, Y., Torensma, B., De Wit, K., Francke, A. L., Oosterveld-Vlug, M. G., & Verheij, R. A. (2023). Opt-in and opt-out consent procedures for the reuse of routinely recorded health Data in scientific research and their consequences for consent rate and consent bias: Systematic review. *Journal of Medical Internet Research*, 25, Article e42131. <https://doi.org/10.2196/42131>
- Deruelle, T., Kalouguina, V., Trein, P., & Wagner, J. (2023). Designing privacy in personalized health: An empirical analysis. *Big Data & Society*, 10(1), 205395172311586. <https://doi.org/10.1177/20539517231158636>
- Eifler, S., & Petzold, K. (2019). Validity aspects of vignette experiments: Expected “what-if” differences between reports of behavioral intentions and actual behavior. In P. Lavrakas, M. Traugott, C. Kennedy, A. Holbrook, E. De Leeuw, & B. West (Eds.), *Experimental methods in survey research* (1st ed., pp. 393–416). Wiley. <https://doi.org/10.1002/9781119083771.ch20>
- Elliott, M. R., & Valliant, R. (2017). Inference for nonprobability samples. *Statistical Science*, 32(2), 249–264. <https://doi.org/10.1214/16-STS598>
- Engström, E., Eriksson, K., Björnstjerna, M., & Strimling, P. (2023). Global variations in online privacy concerns across 57 countries. *Computers in Human Behavior Reports*, 9, 100268. <https://doi.org/10.1016/j.chbr.2023.100268>
- ESS Round 9: European Social Survey (2021). *ESS-9 2018 documentation report. Edition 3.1*. Bergen, European Social Survey Data Archive, NSD–Norwegian Centre for Research Data for ESS ERIC. <https://doi.org/10.21338/NSD-ESS9-2018>
- European Commission. (2023). *A European strategy for data*. <https://digital-strategy.ec.europa.eu/en/policies/strategy-data>
- Garnier, S., Ross, N., BoB Rudis Filipovic-Pierucci, A., Galili, T., Timelyportfolio Greenwell, B., Sievert, C., Harris, D. J., & Chen, J. J. (2021). *sjmgarnier/viridis: Viridis 0.6.0* [pre-CRAN release] (v0.6.0pre) [Computer software]. Zenodo. <https://doi.org/10.5281/ZENODO.4679424>
- Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, 77, 226–261. <https://doi.org/10.1016/j.cose.2018.04.002>
- Gerdon, F., Nissenbaum, H., Bach, R. L., Kreuter, F., & Zins, S. (2021). Individual acceptance of using health data for private and public benefit: Changes during the COVID-19 pandemic. *Harvard Data Science Review* (Special issue 1), 1–24. <https://doi.org/10.1162/99608f92.edf2fc97>
- Ghaiumy Anaraky, R., Li, Y., & Knijnenburg, B. (2021). Difficulties of measuring culture in privacy studies. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–26. <https://doi.org/10.1145/3479522>
- Gilbert, S., Shilton, K., & Vitak, J. (2023). When research is the context: Cross-platform user expectations for social media data reuse. *Big Data & Society*, 10(1), 205395172311641. <https://doi.org/10.1177/20539517231164108>
- Gilbert, S., Vitak, J., & Shilton, K. (2021). Measuring Americans’ comfort with research uses of their social media data. *Social Media + Society*, 7(3), 205630512110338. <https://doi.org/10.1177/20563051211033824>
- Goetzen, A., Dooley, S., & Redmiles, E. M. (2022). Ctrl-Shift: How privacy sentiment changed from 2019 to 2021. *Proceedings on Privacy Enhancing Technologies*, 2022(4), 457–485. <https://doi.org/10.56553/popets-2022-0118>
- Haug, S., Schnell, R., Raptis, G., Dotter, C., & Weber, K. (2023). Wissen und Einstellung zur Speicherung und Freigabe von Gesundheitsdaten. Ergebnisse einer Bevölkerungsbefragung [Knowledge and attitudes toward the storage and release of health data. Results of a population survey]. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen*, 184, 50–58. <https://doi.org/10.1016/j.zefq.2023.11.001>

- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2–3), 61–83. <https://doi.org/10.1017/S0140525X0999152X>
- Hofstede, G. H., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind. Intercultural cooperation and its importance for survival* (3rd ed.). McGraw-Hill.
- Hofstede Insights. (2023). *Country comparison tool*. <https://www.hofstede-insights.com/country-comparison-tool>
- Horne, C., Darras, B., Bean, E., Srivastava, A., & Frickel, S. (2015). Privacy, technology, and norms: The case of Smart Meters. *Social Science Research*, 51, 64–76. <https://doi.org/10.1016/j.ssresearch.2014.12.003>
- Hutchings, E., Loomes, M., Butow, P., & Boyle, F. M. (2020). A systematic literature review of health consumer attitudes towards secondary use and sharing of health administrative and clinical trial data: A focus on privacy, trust, and transparency. *Systematic Reviews*, 9(1), 235. <https://doi.org/10.1186/s13643-020-01481-9>
- Jackson, T. (2020). The legacy of Geert Hofstede. *International Journal of Cross Cultural Management*, 20(1), 3–6. <https://doi.org/10.1177/1470595820915088>
- Jamieson, K. H., Lupia, A., Amaya, A., Brady, H. E., Bautista, R., Clinton, J. D., Dever, J. A., Dutwin, D., Goroff, D. L., Hillygus, D. S., Kennedy, C., Langer, G., Lapinski, J. S., Link, M., Philpot, T., Prewitt, K., Rivers, D., Vavreck, L., Wilson, D. C., & McNutt, M. K. (2023). Protecting the integrity of survey research. *PNAS Nexus*, 2(3), pgad049. <https://doi.org/10.1093/pnasnexus/pgad049>
- Kalkman, S., van Delden, J., Banerjee, A., Tyl, B., Mostert, M., & van Thiel, G. (2022). Patients' and public views and attitudes towards the sharing of health data for research: A narrative review of the empirical evidence. *Journal of Medical Ethics*, 48(1), 3–13. <https://doi.org/10.1136/medethics-2019-105651>
- Kao, Y.-H., & Sapp, S. G. (2022). The effect of cultural values and institutional trust on public perceptions of government use of network surveillance. *Technology in Society*, 70, 102047. <https://doi.org/10.1016/j.techsoc.2022.102047>
- Kassambara, A. (2020). *Ggpubr: "ggplot2" based publication ready plots. R package version 0.4.0*. <https://CRAN.R-project.org/package=ggpubr>
- 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>
- Kemmelmeier, M. (2016). Cultural differences in survey responding: Issues and insights in the study of response biases. *International Journal of Psychology*, 51(6), 439–444. <https://doi.org/10.1002/ijop.12386>
- Kho, M. E., Duffett, M., Willison, D. J., Cook, D. J., & Brouwers, M. C. (2009). Written informed consent and selection bias in observational studies using medical records: Systematic review. *British Medical Journal*, 338, Article b866. <https://doi.org/10.1136/bmj.b866>
- Kim, K. K., Joseph, J. G., & Ohno-Machado, L. (2015). Comparison of consumers' views on electronic data sharing for healthcare and research. *Journal of the American Medical Informatics Association*, 22(4), 821–830. <https://doi.org/10.1093/jamia/ocv014>
- Kim, Y., & Stanton, J. M. (2016). Institutional and individual factors affecting scientists' data-sharing behaviors: A multilevel analysis. *Journal of the Association for Information Science and Technology*, 67(4), 776–799. <https://doi.org/10.1002/asi.23424>
- Kohler, U., & Post, J. C. (2023). Welcher Zweck heiligt die Mittel? Bemerkungen zur Repräsentativitätsdebatte in der Meinungsforschung [Which end justifies the means? Remarks on the debate around representativity in opinion research]. *Zeitschrift für Soziologie*, 52(1), 67–78. <https://doi.org/10.1515/zfsocz-2023-2001>
- Li, Y., Kobsa, A., Knijnenburg, B. P., & Carolyn Nguyen, M.-H. (2017). Cross-cultural privacy prediction. *Proceedings on Privacy Enhancing Technologies*, 2017(2), 113–132. <https://doi.org/10.1515/popets-2017-0019>
- Liu, J. (2022). Social data governance: Towards a definition and model. *Big Data & Society*, 9(2), 205395172211113. <https://doi.org/10.1177/20539517221111352>
- Lüdecke, D. (2023). *sjPlot: Data visualization for statistics in social science. R package version 2.8.14*. <https://CRAN.R-project.org/package=sjPlot>
- Martin, K., & Nissenbaum, H. (2017). Measuring privacy: An empirical test using context to expose confounding variables. *The Columbia Science & Technology Law Review*, 18, 176–218. <https://doi.org/10.7916/STLR.V18I1.4015>
- Masur, P. K., Epstein, D., Quinn, K., Wilhelm, C., Baruh, L., & Lutz, C. (2024). Comparative Privacy Research: Literature review, framework, and research agenda. *SocArXiv*. <https://doi.org/10.31235/osf.io/fjqhs>
- Mood, C. (2010). Logistic regression: Why we cannot do what we think we can do, and what we can do about it. *European Sociological Review*, 26(1), 67–82. <https://doi.org/10.1093/esr/jcp006>
- Mothersbaugh, D. L., Foxx, W. K., Beatty, S. E., & Wang, S. (2012). Disclosure antecedents in an online service context: The role of sensitivity of information. *Journal of Service Research*, 15(1), 76–98. <https://doi.org/10.1177/1094670511424924>
- National Data Guardian for Health and Social Care in England. (2022). *What do we mean by public benefit? Evaluating public benefit when health and adult social care data is used for purposes beyond individual care*. <https://www.gov.uk/government/publications/what-do-we-mean-by-public-benefit-evaluating-public-benefit-when-health-and-adult-social-care-data-is-used-for-purposes-beyond-individual-care/what-do-we-mean-by-public-benefit-evaluating-public-benefit-when-health-and-adult-social-care-data-is-used-for-purposes-beyond-individual-care>
- Newlands, G., Lutz, C., Tamò-Larrieux, A., Villaronga, E. F., Harasgama, R., & Scheitlin, G. (2020). Innovation under pressure: Implications for data privacy during the COVID-19 pandemic. *Big Data & Society*, 7(2), 205395172097668. <https://doi.org/10.1177/2053951720976680>
- Nießen, D., Beierlein, C., Rammstedt, B., & Lechner, C. M. (2021). *Interpersonal trust short scale (KUSIV3)*. ZIS-The Collection of Items and Scales for the Social Sciences. https://doi.org/10.6102/ZIS292_EXZ
- Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press.
- Nissenbaum, H. (2018). Respecting context to protect privacy: Why meaning matters. *Science and Engineering Ethics*, 24(3), 831–852. <https://doi.org/10.1007/s11948-015-9674-9>
- Office of the Australian Information Commission, & Lonergan Research. (2020). *Australian community attitudes to privacy survey 2020*. https://www.oaic.gov.au/__data/assets/pdf_file/0015/2373/australian-community-attitudes-to-privacy-survey-2020.pdf

- Park, Y. J. (2013). Digital literacy and privacy behavior online. *Communication Research*, 40(2), 215–236. <https://doi.org/10.1177/0093650211418338>
- Pew Research Center. (2014). *Pew Research Center's internet project / GfK privacy panel. Survey #1 topline*. https://www.pewresearch.org/wp-content/uploads/sites/9/2015/07/PrivacyPanelTopline_11-12-14.pdf
- Proferes, N., & Walker, S. (2020). *Researcher views and practices around informing, getting consent, and sharing research outputs with social media users when using their public data*. 53rd Hawaii International Conference on System Sciences, Wailea, HI, United States. https://aisel.aisnet.org/hicss-53/dsm/critical_and_ethical_studies/2/
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Roßmann, J. (2010). *Data quality in web surveys of the German longitudinal election study 2009* [Conference presentation]. 3rd ECPR Graduate Conference, Dublin, Ireland.
- Schaffel, N., Schmidt, I., Peiffer, H., & Ellwart, T. (2021). *ICT Self-Concept scale (ICT-SC25)*. ZIS-The Collection of Items and Scales for the Social Sciences. https://doi.org/10.6102/ZIS308_EXZ
- Schomakers, E.-M., Lidynia, C., Müllmann, D., & Ziefle, M. (2019). Internet users' perceptions of information sensitivity—insights from Germany. *International Journal of Information Management*, 46, 142–150. <https://doi.org/10.1016/j.ijinfomgt.2018.11.018>
- Schumacher, C., Eggers, F., Verhoef, P. C., & Maas, P. (2023). The effects of cultural differences on consumers' willingness to share personal information. *Journal of Interactive Marketing*, 58(1), 72–89. <https://doi.org/10.1177/10949968221136555>
- Shaw, J. A., Sethi, N., & Cassel, C. K. (2020). Social license for the use of big data in the COVID-19 era. *npj Digital Medicine*, 3(1), 128. <https://doi.org/10.1038/s41746-020-00342-y>
- Silber, H., Gerdon, F., Bach, R., Kern, C., Keusch, F., & Kreuter, F. (2022). A preregistered vignette experiment on determinants of health data sharing behavior: Willingness to donate sensor data, medical records, and biomarkers. *Politics and the Life Sciences*, 41(2), 161–181. <https://doi.org/10.1017/pls.2022.15>
- Smith, J., Dinev, T., & Xu, H. (2011). Information privacy research: An interdisciplinary review. *MIS Quarterly*, 35(4), 989–1015. <https://doi.org/10.2307/41409970>
- SOEP-IS Group. (2021). *SOEP-IS 2018—Questionnaire for the SOEP innovation sample (Update release 2019)* (SOEP survey papers 948: Series A—Survey instruments [Erhebungsinstrumente]). DIW Berlin/SOEP. https://www.diw.de/documents/publikationen/73/diw_01.c.821936.de/diw_ssp0948.pdf
- Stockdale, J., Cassell, J., & Ford, E. (2019). “Giving something back”: A systematic review and ethical enquiry into public views on the use of patient data for research in the United Kingdom and the Republic of Ireland. *Wellcome Open Research*, 3(6), 1–26. <https://doi.org/10.12688/wellcomeopenres.13531.2>
- Terpstra, A., De Rooij, A., & Schouten, A. (2023). Online proctoring: Privacy invasion or study alleviation? Discovering acceptability using contextual integrity. In *CHI '23: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (p. 167). Association for Computing Machinery. <https://doi.org/10.1145/3544548.3581181>
- Trein, P., & Varone, F. (2023). Citizens' agreement to share personal data for public policies: Trust and issue importance. *Journal of European Public Policy*, 31, 2483–2508. <https://doi.org/10.1080/13501763.2023.2205434>
- Trepte, S. (2020). *The privacy longitudinal study* (2.0.0) [Dataset]. GESIS Data Archive. <https://doi.org/10.7802/2117>
- Trepte, S., & Masur, P. K. (2016). *Cultural differences in media use, privacy, and self-disclosure. Research report on a multicultural survey study*. University of Hohenheim. <https://research.vu.nl/en/publications/cultural-differences-in-social-media-use-privacy-and-self-disclosure>
- Trepte, S., Reinecke, L., Ellison, N. B., Quiring, O., Yao, M. Z., & Ziegele, M. (2017). A cross-cultural perspective on the privacy calculus. *Social Media + Society*, 3(1), 205630511668803. <https://doi.org/10.1177/2056305116688035>
- Utz, C., Becker, S., Schnitzler, T., Farke, F. M., Herbert, F., Schaewitz, L., Degeling, M., & Dürmuth, M. (2021). Apps against the spread: Privacy implications and user acceptance of COVID-19-related smartphone apps on three continents. In *CHI '21: Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (p. 70). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445517>
- Vitak, J., & Zimmer, M. (2020). More than just privacy: Using contextual integrity to evaluate the long-term risks from COVID-19 surveillance technologies. *Social Media + Society*, 6(3), 205630512094825. <https://doi.org/10.1177/2056305120948250>
- Walsh, T. (2023). Modeling COVID-19 with big mobility data: Surveillance and reaffirming the people in the data. *Big Data & Society*, 10(1), 205395172311641. <https://doi.org/10.1177/20539517231164115>
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., . . . Yutani, H. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/10.21105/joss.01686>
- Wnuk, A., Oleksy, T., & Domaradzka, A. (2021). Prosociality and endorsement of liberty: Communal and individual predictors of attitudes towards surveillance technologies. *Computers in Human Behavior*, 125, 106938. <https://doi.org/10.1016/j.chb.2021.106938>

Author Biography

Frederic Gerdon is a postdoc at the Mannheim Center for European Social Research (MZES) at the University of Mannheim, Germany. He uses quantitative methods to study social aspects of digitalization, particularly privacy challenges and algorithmic decision-making.