methods, data, analyses | 2025, pp. 1-42 DOI: 10.12758/mda.2025.05

From Clicks to Quality: Assessing Advertisement Design's Impact on Social Media Survey Response Quality

Jessica Donzowa[1, 2], Simon Kühne[2] & Zaza Zindel[2, 3]

[1] *Max Planck Institute for Demographic Research*
[2] *Bielefeld University*
[3] *German Center for Integration and Migration Research (DeZIM)*

Abstract

Researchers are increasingly using social media platforms for survey recruitment. However, empirical evidence remains sparse on how the content and design characteristics of advertisements used for recruitment affect response quality in surveys. Building on leverage-salience and self-determination theory, we assess the effects of advertisement design on response quality. We argue that different advertisement designs may resonate with specific social groups who vary in their commitment to the survey, resulting in differences in the observed response quality. We use data from a study conducted via ads placed on Facebook in Germany and the United States in June 2023. The survey, focusing on attitudes toward climate change and immigration, featured images with varying thematic associations with the topics (strong, loose, neutral). The sample consisted of 4,170 respondents in Germany and 5,469 respondents in the United States. We compare several data quality indicators, including break-off rate, completion time, non-differentiation, item non-response, passing an attention check question, and follow-up availability, across different advertisement features. Regression analyses indicate differences in response quality across advertisement designs, with a strong thematic design generally being associated with poorer response quality. Strongly themed ad designs are generally associated with higher attrition, non-differentiation, and item non-response, and with a lower probability of passing an attention check and providing an e-mail address for future survey inquiries. Our study advances the literature by highlighting the substantial impact of advertisement design on survey data quality, and emphasizing the importance of tailored decision-making in recruitment design for social media-based survey research.

Keywords: social media recruitment, advertisement design, online survey, survey topic interest, response quality, survey invitation design

The use of social media for (online) survey recruitment has grown over the past decade, with the majority of researchers using Facebook by Meta Inc. as a recruitment tool (Zindel, 2023). Although initially designed for business purposes, research has demonstrated that Meta's advertisement manager is effective for recruiting online survey participants (Grow et al., 2022; Iannelli et al., 2020; Kühne & Zindel, 2020; Pötzschke & Braun, 2017). Similar to companies that use advertisements to promote their services and products, researchers can use advertisements—an image or a video with some text and a link—to recruit respondents. Thus, ads on social media represent a form of digital survey invitation that, unlike email or postal invitations, centers around images or videos to draw users' attention on multi-media platforms. At the same time, concerns about data quality remain, particularly regarding representation (e.g., self-selection biases and non-representativeness) and measurement error (e.g., satisficing behavior due to the lack of interviewer presence; De Man et al., 2021; Heerwegh & Loosveldt, 2008).

Previous studies examining the design effects of ad-based survey invitations in social media have largely focused on response and break-off rates (e.g., Choi et al., 2017, Stern et al., 2022). However, there is almost no empirical evidence on how advertisement properties affect the response quality in Facebook-recruited surveys beyond participation. This is somewhat surprising given the vast literature on the design effects (of invitations) on response quality for mail or web surveys (e.g., Haer & Meidert, 2013; Kaplowitz et al., 2012; Keusch, 2013; Mavletova et al., 2014).

Data quality assessments in the social sciences remain fragmented, and there is a need for systematic frameworks to assess different dimensions of data quality (Birkenmaier et al. 2024). This paper aims to contribute to the broader discussion of data quality by focusing specifically on the intrinsic requirements of social media-recruited survey data, particularly the risk of measurement error. We designed a study that varied images in advertisements used for survey recruitment on Facebook in Germany and the United States in 2023. In addition to a "neutral" set of images, we tested images with varying degrees of association with two survey topics: immigration and climate change. In the analyses, we estimated the effects of these topics and ad image properties on several data

Acknowledgements

Jessica Donzowa gratefully acknowledges the resources provided by the International Max Planck Research School for Population, Health and Data Science (IMPRS-PHDS). This study was funded with support from the Max Planck Institute for Demographic Research, which is part of the Max Planck Society.

Direct correspondence to

Jessica Donzowa, Max Planck Institute for Demographic Research (MPIDR), Rostock, Germany & Bielefeld University, Bielefeld, Germany
E-mail: donzowa@demogr.mpg.de

quality indicators, including survey break-off rate, speeding behavior, non-differentiation, item non-response, attentiveness, and willingness to participate in future surveys. Our approach contributes to the current state of research by a) implementing a study design that deliberately varies ad image properties across survey topics, b) focusing on the general online population in two countries (Germany and the United States), rather than on specific sub-populations, and c) testing a large set of data quality indicators. Thus, this paper is the first comprehensive study of the effects of advertisement design on response quality in surveys recruited via social media.

Background and State of Research

The use of social media—mostly Facebook—for (online) survey recruitment has steadily increased over the past decade (Zindel, 2023). While this method extends the reach of surveys, the methodological implications and the potential biases of social media recruitment are not fully understood (Lehdonvirta et al., 2021). Known biases include skewed sample compositions that favor certain populations, which can affect the reliability and validity of the resulting data (Neundorf & Öztürk, 2023). However, beyond the sample composition, the quality of the responses provided also has an impact on study outcomes.

Figure 1 Advertisement used to recruit respondents via Facebook in the United States. Desktop view.

Visual advertisement design is crucial for social media recruitment (Kühne & Zindel, 2020; Neundorf & Öztürk, 2022). Advertisements on these platforms often rely on visual elements, such as images (rarely videos), which typically make up the majority of an ad's display (see Figure 1). Because the number of texts is limited to just a few lines, the visual components often capture the initial attention of potential survey respondents and establish an initial point of engagement. In the case of image-centric ads—the most common approach in social media recruitment—a key decision regarding about what to display in an image (or multiple images) needs to be made by the researcher. Naturally, the question arises of whether the survey topic is supposed to be reflected in the images, and if so, to what extent. Alternatively, researchers have used neutral images that reflect surveys or public opinion more generally (e.g., by displaying a question mark or speech bubbles).

Existing survey methodological theories and frameworks point to several potential mechanisms through which advertisement design—here: the extent to which a survey topic is displayed in an image—can affect response quality.

The impact of survey recruitment materials—such as the design of a social media ad—on respondents' participation decision and commitment levels can be conceptualized based on the leverage-salience theory (Groves, 1992). Leverage refers to the importance of a feature of an advertisement, such as the image or the topic presented to a potential respondent. Salience, on the other hand, refers to how noticeable or prominent that feature is during the survey invitation process. Images, being the most prominent part of an ad, generally have a high degree of salience. In this context, a user's likelihood of participating in the survey is influenced by the perceived benefits of participation, and by how salient those benefits are made in the ad. Whether a feature of an ad is perceived as beneficial varies by individual characteristics, since people may perceive the same factors as more or less beneficial. Some images may resonate more with specific social groups, such as men, people with lower levels of education or specific political interests (Neundorf & Öztürk, 2022). Since different social groups are expected to show varying levels of interest in the ad, a given ad may appeal to potential respondents differently depending on their cognitive abilities, and their willingness to conscientiously participate in an online survey (Zillmann et al., 2014). Therefore, we argue that different ad designs may not only result in different sample compositions (e.g., in terms of socio-demographics), but also in different response behaviors, and, consequently, in varying levels of data quality.

Beyond socio-demographics and cognitive skills, different ad images can influence the sample composition and data quality with respect to individual motivation. Self-determination theory (Wenemark et al., 2011) not only conceptualizes the decision to participate in a survey, but also distinguishes different levels of commitment to the survey. Self-determination theory distinguishes between autonomous extrinsic motivation, which suggests that participants are motivated by contributing to societal knowledge, and intrinsic motivation, which implies that participants find the task itself such as the survey activity enjoyable. Higher commitment, which is associated with the desire to perform the task well, is generally associated with higher response quality (Wenemark et al., 2011). Because levels and types of commitment have been associated with response quality, ad images are expected to affect the responses' quality by appealing to different motivational types through varying ad content. For example, ad images that clearly reveal that the survey topic is on a highly discussed issue, such as climate change, may systematically attract individuals with autonomous extrinsic motivation, who are motivated to contribute to climate change research. In contrast, neutral images may be more likely to attract intrinsically motivated individuals who enjoy the task of responding to surveys in general, regardless of the specific topic. In summary, leverage-salience theory and self-determination theory suggest that the characteristics and the content of ad images may not only affect the sample characteristics in terms of socio-demographics, but also systematically affect respondents' commitment, motivation, and conscientiousness in completing an online survey. Therefore, the design of ad images is expected to impact response quality.

Previous research on online surveys has shown that interest in the survey topic and the perceived burden of participation influence response quality, with greater interest in the topic being associated with higher response quality (Galesic, 2006). Conversely, an increasing desire to abandon the survey due to the response burden is reflected in decreased quality, which is particularly evident just before respondents drop out through increased item non-response (Galesic, 2006). This behavior, which is also referred to as "satisficing" or "short-cutting," refers to the tendency of respondents to choose easier response strategies to minimize their effort and individual survey burden, which may compromise data quality (Krosnick, 1999). Satisficing can include behaviors such as choosing no response (i.e., item non-response), repeating the same answers across various questions (i.e., non-differentiation or "straightlining"), and consistently agreeing with survey items (i.e., acquiescence or "yes-saying"). Research suggests that satisficing behavior is particularly likely to occur when the task difficulty is high and the respondent motivation is low (Holbrook et al., 2003; Kaminska et al., 2010; Roberts et al., 2019).

Despite the prevalence of social media recruitment, our understanding of how advertisement design influences response quality remains limited. Choi et al. (2017) found that the choice of image and wording had an impact on men's engagement levels and time spent on a mental health survey. Similarly, Stern et al. (2022) found that among young men from sexual minorities in the United States, advertisements with images resulted in fewer non-substantive survey responses than ads that used video as visual element, although the response rates were comparable. Neundorf and Öztürk (2022) examined the effects of incentive-based versus thematic advertisements in Turkey. They found that the differences in attentiveness (passing an attention check question by checking the option "Do not know") disappeared when controlling for demographic characteristics. However, while the respondents recruited through incentive-based advertisements were more likely to answer to open-ended questions, they left shorter responses than the participants recruited through the thematic advertisements. Finally, the participants recruited through incentive-based advertisements were more likely to participate in a follow-up survey when contacted again (Neundorf & Öztürk, 2022). Donzowa et al., (2023) showed that during the early stages of the COVID-19 pandemic, more explicit survey topic display was associated with higher numbers of link clicks and higher completion rates. These results suggest that while there are economic advantages to making survey topics more explicit, the impact on response quality remains uncertain.

The theories outlined above describe the pathway through which ad design affects response quality. However, the direction of this effect is still unknown. Following the leverage-salience theory, one could argue that "thematic" advertisements with a more explicit topic presentation tend to recruit more thematically motivated respondents with a strong commitment to the survey topic, resulting in higher response quality. On the other hand, according to self-determination theory, "neutral" advertisements that make no reference to a specific survey topic may recruit respondents who are more intrinsically motivated to respond to surveys in general, resulting in consistently higher response quality. Existing research has reported inconsistent findings regarding the direction and the magnitude of this effect. Our study aims to contribute to the current state of research by implementing a design that allows us to examine the effects of ad design properties on response quality in a general online population survey setting in two countries.

Data & Methods

The following chapter presents the study design used in this project. We also present the indicators used to measure response quality and explain how they are defined.

Study Design

We use data from a survey conducted via Facebook in both Germany (June 25, 2023, to July 2, 2023) and the United States (June 25, 2023, to July 3, 2023). The survey focused on the subjects of climate change and immigration—two topics that are the subject of intense media and public debate in both countries. The advertisement campaign used a variety of images with different thematic associations, as shown in Appendix Figure A1. Fifteen images were used in each country. Thirteen images were the same for both countries, and two images were adapted for each country. For an image showing a flag, an image of the European flag was used in Germany, and an image of the US flag was used in the United States. In the second case, one of the neutral images, images with text in the local language of each country were used ("Your opinion matters" in the United States and "Ihre Meinung ist uns wichtig" in Germany). Images were selected through a multi-step selection process. First, a broad set of images was retrieved from stock image websites (AdobeStock (https://stock.adobe.com) and iStock (https://www.istockphoto.com)), which the research team narrowed down to 73 images (43 for climate change and 30 for immigration). In a second step, these images were evaluated for their association with the survey themes through an image selection survey among the researchers' scientific network. This resulted in a set of 34 images with accompanying information about the association to the survey topic, that is, strong, loose, or neutral. In a third step, these images were used in a survey pretest based on social media recruitment via Facebook and then evaluated for their recruiting performance. The top three images for each association were selected for the final data collection.

Technically, we implemented each image in an individual campaign on Meta's advertisement management system, resulting in a total of 30 campaigns for both countries combined. The campaign's objective was to drive traffic, with the optimization goal of "link-clicks." Meta Pixel was not used in this study. From a research ethics perspective, the use of Meta Pixel may be viewed critically as it provides Meta with information about respondent behavior outside of the Facebook platform to optimize survey completion. The remaining advertising options, such as platform placement (Facebook Newsfeed) or sender of the study, were kept consistent across all ads to ensure that all campaigns had an equal chance to capture the attention of social media users and could be effectively compared. Advertisements ran exclusively on Facebook and did not include Instagram. Only demographic targeting tools were used. Geographic targeting included the respective countries, that is, regions of the U.S. and regions of Germany. The age range for all ads was set to include users of age 18 and older. We tracked through which ad a user entered our survey through Meta Ad Manager's built-in ability to define URL parameters.

When Facebook users were exposed to the ads, they had the option to self-select into the survey by clicking on the ad. Upon clicking, they were redirected to the web survey, which was hosted on Bielefeld University server and implemented using the LimeSurvey software. The web survey was optimized for mobile devices to ensure functionality and proper design across a wide range of computer and mobile device hardware and software. Prior to beginning the survey, participants were informed of the estimated length of the survey and asked for their consent. The survey included questions about the participants' general political interests, as well as their views on immigration and climate change. Respondents were randomly assigned to begin with questions related to either immigration or climate change, irrespective of the ad image they initially encountered.

In addition to the social media recruitment efforts, the survey was replicated using a commercial panel company to allow for comparisons with other online survey populations. However, this comparison should be interpreted with caution, as the terms of participation for this survey differed from those for our social media sample, particularly in terms of compensation, as online panel respondents may be incentivized to participate. For more details on the online access panel data, see Section A2 in the appendix.

Response Quality Measures

We rely on several data quality indicators that are regularly used in the literature (i.e., survey break-off rate, speeding, non-differentiation, item non-response, passing an attention check question, and willingness to participate in future surveys). As a predictor of potential differences in these quality indicators, we use the advertisement design through which the respondent entered the survey.

Survey break-off rate is an important measure of data quality, because it indicates the respondents' overall motivation to respond to the whole questionnaire (Tangmanee & Niruttinanon, 2019). Previous research has shown how this rate can be influenced by survey design features, such as including a progress bar or announcing the survey length (Liu et al., 2016). We calculate the break-off rate as the ratio of the number of surveys started—defined as respondents proceeding past the welcome page, thereby agreeing to participate and consenting to the processing of personal data—to the total number of surveys. This is calculated separately for each advertisement design. In the regression analysis, P refers to the probability of having started but not completed the survey, as opposed to having completed the survey.

Next, we define the outcome of *speeding*. Providing answers quickly usually indicates that a respondent wants to finish the survey without giving enough thought to the questions to provide accurate answers (Zhang & Conrad, 2014). However, the interpretation of a short response time is not straightforward, as it could also indicate that the respondents have stable and crystallized opinions about certain topics, or that the survey design is efficient (Zhang & Conrad, 2014). Nevertheless, survey completion time is a commonly used data quality indicator that reflects possible general problems with the survey itself or the motivation of (some) respondents to answer the questions thoroughly. The survey completion times presented here are calculated based only on completed interviews—that is, surveys that reached the final page of the web survey, regardless of item non-response. In the multivariate analysis, we transform the completion time into a binary variable, defining speeding as having a completion time in the fastest 10% of the sample distribution. This means that the completion time is less than 9.94 minutes for Germany and less than 10.60 minutes for the United States. P refers to the probability of speeding.

We analyze *non-differentiation* in the context of satisficing behavior. This behavior may result from a lack of motivation or response-ability (Gao et al., 2016; Roberts et al., 2019). We use a battery of eight items that measure attitudes toward immigration and estimate the number of inconsistent responses. The response scale ranged from fully agree to fully disagree on a five-point scale and included an option for no opinion, which was excluded from this analysis. For the immigration items, half of the statements were framed with positive attitudes toward immigration, and the other half were framed with negative attitudes toward immigration (see Appendix Table A3). We assume that an attentive respondent would tend to agree with half of the items and tend to disagree with the other half. In order to assess the consistency of response behavior, the rating scales of the items were re-coded to point in the same direction. In the following steps, the mean value was calculated for the group of positively and negatively framed statements. Next, the absolute difference between these means was calculated. If the responding behavior was consistent, the difference should be close to zero, while higher values should indicate inconsistent responding behavior. When respondents reached the maximum value of four, this means that they fully agreed with one framing and also fully agreed with items with the contradictory framing, indicating inconsistent response behavior. For the regression analysis, we categorize this outcome into two categories: low non-differentiation (0) is assigned for values below the median value of 0.5; while high non-differentiation (1) is assigned to all values above the threshold of the median value of 0.5. P refers to the probability of high non-differentiation.

Item non-response means that participants started to answer the questionnaire, but did not answer certain questions where a response would have been expected (Cehovin et al., 2023). Respondents may choose not to answer a particular question for many reasons. These include not knowing or remembering the answer, privacy concerns, or a lack of motivation. In this regard, research has shown that adding motivational statements after a question is left unanswered reduces item non-response in self-administered surveys (Al Baghal & Lynn, 2015). In our study, item non-response is defined as seeing a survey question but not responding to it. There were no compulsory questions in the survey. Providing non-substantive answers (e.g., “prefer not to say” or “other”) does not count as non-response. Respondents who did not start the survey are excluded. The percentage of non-response to the survey questions was calculated as the percentage of missing responses (i.e., a question that was seen but not answered) divided by the number of expected responses, which is the sum of the number of times a valid response was recorded and the number of times the question was seen and no response was recorded.

For the regression analysis, we categorize item non-response into two categories: zero for no item non-response and one if respondents did not respond to at least one item. *P* refers to the probability of item non-response.

Next, we consider *attentiveness*, which is measured by an attention check question in the form of an instructed response item (IRI) developed by Gummer et al., (2018). These items are included as part of a grid of questions in which one item asks respondents to select a particular response category. This assesses whether respondents have read the text of the particular item. Failure to provide the required response indicates inattention due to insufficient reading or understanding of the particular item (Gummer et al., 2018). In our survey, the IRI was administered in a list of six statements about politics and society (see Appendix Table A5 for the question text). Respondents were asked to indicate their opinion regarding these statements on a five-point scale, ranging from "strongly agree" to "strongly disagree." Item four of the six was not a political statement, but instructed respondents to choose a specific value on the response scale ("Please click 'rather disagree'."). From this, we construct our measure of attention by defining the attention check as "passed" (1) if the required category was selected, and as "failed" (0) if any other category was selected. Item non-response to this question was excluded before defining these categories. From this, we calculate the percentage of respondents who passed the attention check for each ad design. For the regression analysis, *P* refers to the probability of passing the attention check.

At the end of the survey, respondents were asked if they would like to provide their e-mail address so they could be contacted for future surveys. Loosveldt and Storms (2008) show that the *willingness to participate in future surveys* is influenced by the respondent's overall opinion of surveys. Willingness to participate in the future is increased when respondents perceive surveys as a useful tool for sharing their opinions. On the other hand, the likelihood of future participation is reduced when respondents perceive the investment of time and cognitive effort required as too high, or when there are concerns about data privacy (Loosveldt & Storms, 2008). Willingness to participate in future surveys is measured by the percentage of respondents who provided an e-mail address. This is calculated as the number of e-mail entries divided by the total; that is, the sum of the entries and the empty entries (i.e., the sum of respondents who saw the question and did not enter an e-mail in the open text field.). For the regression analysis, the binary outcome of providing an e-mail address (1) or not providing an e-mail address (0) is used. *P* refers to the probability of providing an e-mail address.

Regression Analysis

We use logistic regression to estimate the effects of advertisement design on each binary data quality outcome:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 x_{ad_{imm-strong}} + \beta_2 x_{ad_{imm-loose}} + \beta_3 x_{ad_{clim-strong}} + \beta_4 x_{ad_{clim-loose}} + \beta_{5_{sex}} + \beta_6 x_{age_{40-59}} + \beta_7 x_{age_{60-64}} + \beta_8 x_{age_{65+}} + \beta_9 x_{income_{low}} + \beta_{10} x_{income_{median}} + \beta_{11} x_{income_{missing}} + \beta_{12} x_{device}$$

The independent variables in the models above correspond to the following categories:

- *Advertisement design* refers to the ad image through which a Facebook user reached our survey. These are: "immigration-strong" (imm-strong), "immigration-loose" (imm-loose), "climate-strong" (clim-strong), "climate-loose" (clim-loose), and the "neutral" design. We use the "neutral" design as the reference.
- *Sex* refers to the respondents' self-reported sex (i.e., male, female). We use "female" as the reference.
- *Age* represents the respondents' age, grouped into four age categories (i.e., 18–39, 40–59, 60–64, 65+). We use the "18-39" age group as the reference.
- *Income* refers to the monthly household income, with the categories "low" (≤ 25th percentile), "medium" (> 25th percentile and ≤ 50th percentile), "high" (> 50th percentile), and "missing." We use "high" as the reference.
- *Device* refers to the device that the respondents used to fill out the survey. It consists of the categories "mobile" (Android Smartphone/Tablet and iPad/iPhone) and "desktop." We use "desktop" as the reference.

For the regression analysis, the missing values for gender and age were removed. Full case analysis is considered more valid than using an imputation technique that would estimate socio-demographic information such as age and sex based on limited information in the survey. The analysis was conducted using R version 4.3.2. A complete list of the R packages used in the analysis can be found in the Appendix Section A4.

Results

Recruitment Results and Response Quality

In terms of campaign performance, in Germany, the design strongly related to immigration received the most link clicks (1,646) and the neutral design received the fewest clicks (859). In the United States, the design loosely related to climate change received the most clicks (2,778) and the neutral design received the fewest clicks (1,273) (see Appendix Table A1). Using the impression and reach performance metrics provided by Meta (see Appendix Table A1), we calculate an indicator of the average number of times each ad was shown to a user. This shows that, on average, a user had a chance to see our ads between 1.3 to 1.4 times. Since we cannot assume that each user consciously saw the ads for each impression, this measure can be interpreted as an upper bound estimate and the vast majority of respondents most likely saw only one specific ad. In Germany about 890 euros were spent, with a cost per click of between 0.11 and 0.21 euros. Recruitment costs were higher in the United States, with 3,457 euros spent and a cost per click ranging from 0.25 to 0.55 euros.

Initially, 6,827 respondents in Germany and 12,596 in the United States reached our survey website. We had to exclude those cases with no information in the predictor variable of advertisement design (this removes 159 cases for Germany and 190 cases for the United States). The Facebook sample consisted of 6,668 respondents for Germany and 12,406 for the United States, including both started and not started surveys. In Germany, 4,170 respondents started the survey and 2,495 completed it. In the United States, 5,469 respondents started the survey and 2,520 completed it (see Appendix Table A2).

Among the respondents who started surveys in Germany, 31% were recruited through a design strongly related to immigration, while the smallest share (14%) were recruited through a design loosely related to climate. Among respondents who started surveys in the United States, 26% were recruited through the design strongly related to climate, while the smallest shares were recruited through the neutral and the design loosely related to immigration (each 16%) (see Appendix Table A1).

In both countries, the samples consist of about 50% men. Approximately 40% of all respondents are female, and 7–8% of the participants did not provide gender information. On average, the participants were 74 years old in the United States (range: 19–96, *SD* = 10) and were 60 years old in Germany (range: 18–99, *SD* = 11). However, 10% of respondents in Germany and 12% of respondents in the United States did not report their age (see Appendix Table A4).

Table 1 provides a descriptive overview of the quality indicators by country. Only started surveys are included in these results. The survey break-off rate is higher in the United States (54%) than in Germany (40%). On average, the respondents completed the survey in 16 minutes in Germany and 18 minutes in the United States. The rate of non-differentiation is higher in Germany (0.5) than in the United States (0.4). In both countries, there is about 3% item non-response to the survey questions. More participants passed the attention check question in the United States (77%) than in Germany (64%). Finally, the willingness to provide an e-mail address is higher in the United States (53%) than in Germany (42%) (see Table 1).

We also examine the changes in the quality indicators over the eight-day recruitment period by advertisement design. There are no systematic time trends in the evolution of the quality indicators, suggesting that the algorithmic placement of the ads does not promote specific response quality types (see Appendix Figure A2).

Compared to the online panel respondents, the social media sample has a significantly higher break-off rate. However, online panel respondents were incentivized to complete the survey. Average completion time and item non-response rates are lower for the online panel than for the social media sample. The rate of passing the attention check is higher in the online panel. On the other hand, the rate of non-differentiation is higher in the online panel than in the social media sample. See Section A2 in the appendix for a description of the online panel sample.

Table 1 Descriptive statistics by country showing the survey quality indicators: break-off rate, mean completion time, non-differentiation, item non-response, and e-mail provision

Indicator	Germany		United States	
Break-off rate (%)	40.17	[38.69, 41.66]	53.92	[52.60, 55.24]
Mean completion time (min)	16.11	[15.78, 16.43]	17.93	[17.57, 18.30]
Non-differentiation	0.51	[0.47, 0.54]	0.40	[0.37, 0.44]
Item non-response (%)	2.91	[2.85, 2.97]	3.19	[3.13, 3.26]
Attention check passed (%)	63.51	[61.57, 65.40]	76.73	[75.02, 78.35]
Provided e-mail address (%)	42.30	[40.47, 44.15]	52.85	[50.99, 54.70]

Notes: Values in brackets refer to the 95% confidence interval. Unweighted.

Advertisement Design Effects on Data Quality

This chapter presents the results of the binary logistic regression analysis used to assess the six quality indicators separately for Germany and the United States. Figure 2 shows the odds ratio estimates controlling for gender, age, income, and the device used to answer the survey, with the neutral design as the reference category.

We first present the results for Germany. We find a higher probability of leaving the survey (OR = 1.6, 95% CI [1.03, 2.42])[1] for the design classified as strongly related to immigration. In terms of speeding—that is, having a survey completion time in the top 10th percentile—we see that the design classified as strongly related to climate is associated with a lower likelihood of speeding than the neu-

[1] OR = odds ratio, CI = confidence interval.

tral ad design (OR = 0.6, 95% CI [0.41, 0.95]). The design classified as strongly related to immigration is associated with a lower chance of passing the attention check (OR = 0.6, 95% CI [0.43, 0.76]). Finally, respondents recruited by the design classified as having a strong topic relation (both immigration (OR = 0.5, 95% CI [0.39, 0.65]) and climate (OR = 0.6, 95% CI [0.45, 0.75])) are less likely to provide an e-mail address than respondents recruited by the neutral design. There is no design effect on non-differentiation and item non-response (see Figure 2).

For the United States, we see a correlation with a lower likelihood of speeding (OR = 0.5, 95% CI [0.32, 0.75] for climate and OR = 0.6, 95% CI [0.38, 0.88] for immigration) and a higher likelihood of non-differentiation (OR = 1.5, 95% CI [1.16, 2.00] for climate and OR = 1.4, 95% CI [1.08, 1.87] for immigration) for the ads classified as having a strong topic relation compared to the neutral ads. The design classified as strongly related to climate is also associated with a higher probability of item non-response (OR = 1.4, 95% CI [1.04, 1.87]) and a lower probability of passing the attention check (OR = 0.7, 95% CI [0.47, 0.94]). The results for panel availability show that respondents recruited through the design classified as having a strong immigration relation are associated with a lower probability (OR = 0.7, 95% CI [0.56, 0.96]) of providing their e-mail address. The break-off rate and the speeding behavior are not associated with any specific ad design (see Figure 2). It is worth noting that the thematically loosely associated design does not differ significantly from the neutral design in any of the final models for either country.

Full stepwise regression estimates can be found in Appendix Section A3. Looking at the stepwise regression estimates, we can see that for Germany, there is no association between ad design and non-differentiation or item non-response in any of the stepwise models (see Appendix Tables A13 and A15). Similarly, in the United States, survey break-off is not associated with any particular ad design in the separate model steps (see Appendix Table A10). In some cases, initial design effects can be explained by the sample composition. The lower chance of speeding for the designs classified as having a loose topic relation is explained by the age of the respondents (climate) and the device used to fill out the survey (immigration) (see Appendix Table A12). Participation through the design classified as strongly related to immigration is associated with a lower chance of passing the attention check question, this association is explained by the gender of the participants (see Appendix Table A18). For the design classified as loosely related to climate change, the correlation with a lower probability of panel availability is explained by age structure (see Appendix Table A20).

In summary, the designs classified as highly related to the survey topic are associated with lower response quality. In Germany, higher break-off rates, lower odds of speeding, passing the attention check, and panel availability are associated with ad designs classified as strongly related to the survey topic. In the United States, the designs classified as strongly topic-related are associated with a lower likelihood of speeding, a higher likelihood of non-differentiation and item non-response, and lower rates of passing the attention check and providing an e-mail address.

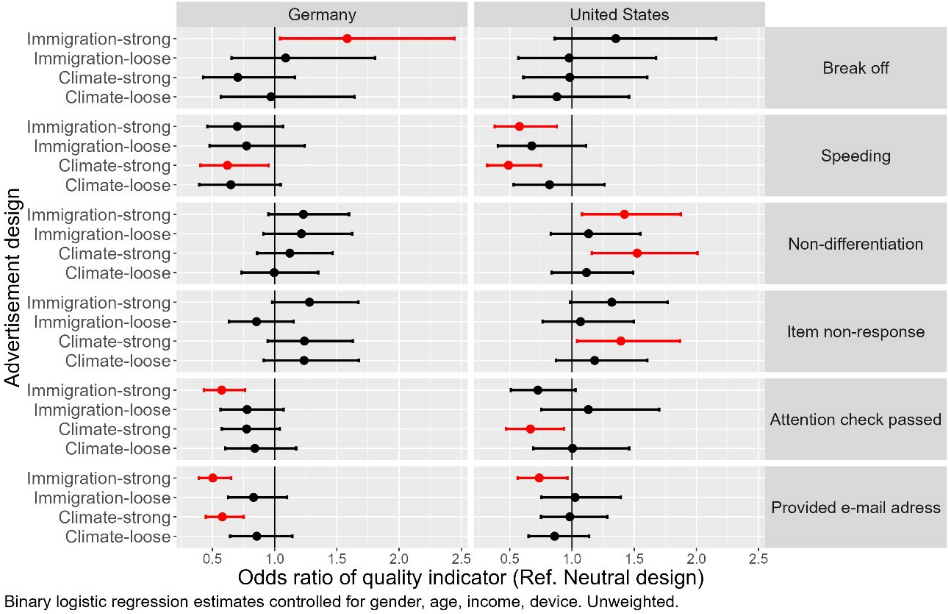


Figure 2 Stepwise binary logistic regression estimates for quality indicators, controlled for advertisement design, gender, age, income, and device for Germany and the United States. Unweighted. Odds ratios significantly different from one are marked in red.

Discussion

In this paper, we examined how advertisement design affects response quality in surveys recruited via social media. Using a study conducted via Facebook in Germany and the United States, we varied the design of the advertisement used for survey recruitment. We hypothesized that a more explicit display of the survey topic would result in systematically different response behavior compared to designs with fewer or no references to the survey topic. However, previous research was inconclusive about the direction of this effect.

We analyzed the impact of ad design on six response quality indicators: survey break-off rate, speeding, non-differentiation, item non-response, passing an attention check, and willingness to participate in a panel.

Consistent with previous findings, we observed that more respondents were recruited through advertisements with a prominent survey topic (i.e., immigration or climate change). However, after controlling for differences in sample composition by gender, age, income, and device, we found that these advertisements were associated with lower response quality. Specifically, ads classified as having a strong survey topic relation were associated with higher break-off rates, longer completion times, more non-differentiation, and more item nonresponse, as well as with lower rates of passing attention checks and willingness to participate in future surveys. We uncovered differences in ad design effects across countries: while ad design had no effect on break-off rates in the United States, it did in Germany; and while ad design had no effect on item nonresponse and non-differentiation in Germany, it did in the United States. We also found that the ads classified as less explicit (loose association) did not differ significantly from the neutral designs in terms of response quality after adjusting for sample composition. While these ads had higher reach and link clicks than neutral ads, they did not lead to higher start rates.

We considered two theoretical perspectives on how ad design might affect response quality. One perspective suggested that a more explicit display of the survey topic would attract highly thematically motivated respondents, thereby improving response quality. The other perspective posited that neutral ads would attract highly intrinsically motivated participants, leading to higher response quality. The results support the second argument, showing that neutral ad designs were associated with higher response quality than ads classified as having strong thematic references. This suggests that higher thematic motivation may lead to the recruitment of respondents with lower levels of response commitment, resulting in more inconsistent response behavior and lower overall response quality. This, in turn, supports our assumption that a more explicit display of the survey topic would lead to systematically different response styles than those expected from respondents recruited through ads with a less salient display of the survey topic. Finally, we want to address two emerging challenges for survey recruitment via social media advertisements: the rise of large language models and AI tools. One emerging challenge for data quality in social media-recruited surveys is the increasing use of large language models and chatbots to fabricate survey responses. These tools can undermine the authenticity and reliability of survey data, introducing new forms of bias and error. Depending on the complexity of the models, conventional survey quality indicators may not be able to distinguish between fabricated and genuine survey responses (Höhne et al. 2024). Future studies should proactively address this issue, as it is critical to maintaining the integrity of web survey-based research.

The use of generative AI tools to create ad images and text in social media recruitment campaigns, may, on the one hand, help to streamline the creative process, allowing for the rapid creation of visually engaging and personalized content, increasing ad reach or engagement. However, it also raises issues of authenticity and audience trust. For example, overly polished or artificial-looking ads may increase user skepticism or reduce perceived credibility, which can impact survey participation rates. In addition, AI tools are trained on existing datasets, which can (re)introduce unintentional biases into ad images or messaging, which could affect the inclusivity and representativeness of survey samples. Future research is needed to explore this systematically. For example, A/B testing could be used to compare AI-generated ads with traditional ads in terms of response rates, participant demographics, and data quality.

Conclusion

Our study shows that advertisement design significantly affects response quality in social media-recruited surveys, with effects varying across different quality indicators and countries. Ads with higher topic salience tend to attract more clicks at a lower cost, but they often result in poorer response quality, including inconsistent responses and higher non-response rates. Conversely, neutral ads tend to yield higher response quality, making them more suitable for general research purposes. The findings have important implications for researchers planning future survey recruitment ad campaigns using Facebook. Specifically, there appears to be a trade-off between the level of attention generated by ads focused on prominent issues such as immigration and climate change (as indicated by higher reach and link clicks) and the quality of survey responses obtained.

While themed ads may initially lower recruitment cost and increase sample size, these benefits can be offset by a higher proportion of low-quality responses. The variation in design effects across countries also highlights the importance of considering country-specific contexts when designing a recruitment campaign that focuses on potentially polarizing social issues. Therefore, the specific objectives of the recruitment campaign should guide the choice of ad design.

However, our study also has several limitations that need to be acknowledged. Because we lacked a direct measure of the level of commitment evoked by the advertisements, we could only assume that higher topic salience correlates with higher thematic motivation.

The classification of images as loosely or strongly related to immigration or climate change is an individual and subjective interpretation and may vary from respondent to respondent. However, we base our image classification on our image selection survey, thus providing empirical support for our classification. Additionally, by excluding cases with missing age or sex information from the regression analysis, we may have underestimated the impact of ad design on response quality by removing the lowest-quality responses. Nevertheless, our models were robust to the inclusion of these missing values.

The generalizability of our findings is also limited by the exclusive use of Facebook for recruitment, which may not translate to other social media platforms such as Instagram, TikTok, or LinkedIn, each having different user demographics and engagement behaviors. In addition, an inherent limitation of studies using any social media platform is the underlying advertising algorithm, which remains a black box to researchers and may change over time, requiring frequent re-evaluation of any methodological finding to ensure robust results. Reliance on commercial platforms for data collection carries additional risks, as platform policies, access, and available features used for data collection, such as the business advertising manager for Meta, may change or be deprecated. This can limit data availability and affect the reproducibility of studies, as has been shown previously for studies using data obtained from social media platforms through API access (Davidson et al., 2023; Freelon, 2018).

In addition, the study focused on two topics that are heavily discussed and politically charged. While such topics are often the focus of social science research projects, our findings are limited, and might not be generalizable to other prominent but less controversial topics. Additionally, the study was conducted in Germany and the United States, where survey recruitment via ads is relatively common. It remains uncertain whether these findings will hold true in countries where this recruitment approach is newer, or where there is greater skepticism towards online ads or invitations.

Future research should address these limitations by exploring the impact of advertisement design across different social media platforms (such as Instagram, X, or TikTok) and a wider range of topics and contexts. By addressing these issues, future studies can build on our findings to further our understanding and optimize the use of social media advertisements to recruit survey participants.

Code and data availability

The data that support the findings of this study are available for scientific purposes upon request at: https://pub.uni-bielefeld.de/record/3002204

The code that was used for the analysis can be found at: osf.io/n76vu

Ethics approval and consent to participate

This study received ethics approval from the Ethics Council of the University of Bielefeld (Application Nr: 2022-209). Electronic informed consent was obtained from all participants who actively opted to participate in the online survey, enabling the collection, storage, and processing of their answers. All participants' data were treated anonymously. Participation was voluntary.

References

Al Baghal, T., & Lynn, P. (2015). Using motivational statements in web-instrument design to reduce item-missing rates in a mixed-mode context. *Public Opinion Quarterly, 79*(2), 568–579. https://doi.org/10.1093/poq/nfv023

Birkenmaier, L., Daikeler, J., Fröhling, L., Gummer, T., Lechner, C., Lux, V., Schwalbach, J., Silber, H., Weiß, B., Weller, K., Wolf, C., Abel, D., Breuer, J., Dietze, S., Dimitrov, D., Döring, H., Hebel, A., Hochman, O., Jünger, S., ... Ziaja, S. (2024). *Defining and evaluating quality for the sciences: Position paper* (GESIS Papers, 2024/06). GESIS – Leibniz-Institut für Sozialwissenschaften. https://doi.org/10.21241/ssoar.96764

Cehovin, G., Bosnjak, M., & Lozar Manfreda, K. (2023). Item nonresponse in web versus other survey modes: A systematic review and meta-analysis. *Social Science Computer Review, 41*(3), 926–945. https://doi.org/10.1177/08944393211056229

Choi, I., Milne, D. N., Glozier, N., Peters, D., Harvey, S. B., & Calvo, R. A. (2017). Using different Facebook advertisements to recruit men for an online mental health study: Engagement and selection bias. *Internet Interventions,* 8, 27–34. https://doi.org/10.1016/j.invent.2017.02.002

Davidson, B. I., Wischerath, D., Racek, D., Parry, D. A., Godwin, E., Hinds, J., Van Der Linden, D., Roscoe, J. F., Ayravainen, L., & Cork, A. G. (2023). Platform-controlled social media APIs threaten open science. *Nature Human Behaviour, 7*(12), 2054–2057. https://doi.org/10.1038/s41562-023-01750-2

De Man, J., Campbell, L., Tabana, H., & Wouters, E. (2021). The pandemic of online research in times of COVID-19. *BMJ Open, 11*(2), e043866. https://doi.org/10.1136/bmjopen-2020-043866

Donzowa, J., Perrotta, D., & Zagheni, E. (2023). *Assessing self-selection biases in online surveys: Evidence from the COVID-19 Health Behavior Survey* (MPIDR Working Paper WP-2023-047). Max Planck Institute for Demographic Research. https://doi.org/10.4054/MPIDR-WP-2023-047

Freelon, D. (2018). Computational research in the post-API age. *Political Communication, 35*(4), 665–668. https://doi.org/10.1080/10584609.2018.1477506

Galesic, M. (2006). Dropouts on the web: Effects of interest and burden experienced during an online Survey. *Journal of Official Statistics, 22*(2), 313–328. https://www.proquest.com/docview/1266792615?pq-origsite=gscholar&fromopenview=true&sourcetype=Scholarly%20Journals

Gao, Z., House, L., & Bi, X. (2016). Impact of satisficing behavior in online surveys on consumer preference and welfare estimates. *Food Policy, 64,* 26–36. https://doi.org/10.1016/j.foodpol.2016.09.001

Groves, R. M., Cialdini, R. B., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *The Public Opinion Quarterly, 56*(4), 475–495. http://www.jstor.org/stable/2749203

Grow, A., Perrotta, D., Del Fava, E., Cimentada, J., Rampazzo, F., Gil-Clavel, S., Zagheni, E., Flores, R. D., Ventura, I., & Weber, I. (2022). Is Facebook's advertising data accurate enough for use in social science research? Insights from a cross-national online survey. *Journal of the Royal Statistical Society Series A: Statistics in Society, 185*(Supplement_2), S343–S363. https://doi.org/10.1111/rssa.12948

Gummer, T., Roßmann, J., & Silber, H. (2018). Using instructed response items as attention checks in web surveys: Properties and implementation. *Sociological Methods & Research, 50*(1), 238-264. https://doi.org/10.1177/0049124118769083

Haer, R., & Meidert, N. (2013). Does the first impression count? Examining the effect of the welcome screen design on the response rate. *Survey Methodology, 39*(2), 419–434. http://www.statcan.gc.ca/pub/12-001-x/2013002/article/11885-eng.htm

Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high internet-coverage population: Differences in response quality. *Public Opinion Quarterly, 72*(5), 836–846. https://www.jstor.org/stable/25548047

Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly, 67*(1), 79–125. https://doi.org/10.1086/346010

Höhne, J. K., Claassen, J., Shahania, S., & Broneske, D. (2024). Bots in web survey interviews: A showcase. *International Journal of Market Research, 67*(1), 3–12. https://doi.org/10.1177/14707853241297009

Iannelli, L., Giglietto, F., Rossi, L., & Zurovac, E. (2020). Facebook digital traces for survey research: Assessing the efficiency and effectiveness of a Facebook ad–based procedure for recruiting online survey respondents in niche and difficult-to-reach populations. *Social Science Computer Review, 38*(4), 462–476. https://doi.org/10.1177/0894439318816638

Kaminska, O., McCutcheon, A. L., & Billiet, J. (2010). Satisficing among reluctant respondents in a cross-national context. *Public Opinion Quarterly, 74*(5), 956–984. https://doi.org/10.1093/poq/nfq062

Kaplowitz, M. D., Lupi, F., Couper, M. P., & Thorp, L. (2012). The effect of invitation design on web survey response rates. *Social Science Computer Review, 30*(3), 339–349. https://doi.org/10.1177/0894439311419084

Keusch, F. (2013). The role of topic interest and topic salience in online panel web surveys. *International Journal of Market Research, 55*(1), 59–80. https://doi.org/10.2501/IJMR-2013-007

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology, 50*(1), 537–567. https://doi.org/10.1146/annurev.psych.50.1.537

Kühne, S., & Zindel, Z. (2020). Using Facebook and Instagram to recruit web survey participants: A step-by-step guide and application. *Survey Methods: Insights from the Field, Special Issue: 'Advancements in Online and Mobile Survey Methods'.* https://doi.org/10.13094/SMIF-2020-00017

Lehdonvirta, V., Oksanen, A., Räsänen, P., & Blank, G. (2021). Social media, web, and panel surveys: Using non-probability samples in social and policy research. *Policy & Internet, 13*(1), 134–155. https://doi.org/10.1002/poi3.238

Liu, M., Kuriakose, N., Cohen, J., & Cho, S. (2016). Impact of web survey invitation design on survey participation, respondents, and survey responses. *Social Science Computer Review, 34*(5), 631–644. https://doi.org/10.1177/0894439315605606

Loosveldt, G., & Storms, V. (2008). Measuring public opinions about surveys. *International Journal of Public Opinion Research, 20*(1), 74–89. https://doi.org/10.1093/ijpor/edn006

Mavletova, A., Deviatko, I., & Maloshonok, N. (2014). Invitation design elements in web surveys – Can one ignore interactions? *Bulletin of Sociological Methodology/Bulletin de Methodologie Sociologique, 123*(1), 68–79. https://doi.org/10.1177/0759106314531883

Neundorf, A., & Öztürk, A. (2022). *Advertising online surveys on social media: How your advertisements affect your samples.* OSF. https://doi.org/10.31219/osf.io/84h3t

Neundorf, A., & Öztürk, A. (2023). How to improve representativeness and cost- effectiveness in samples recruited through meta: A comparison of advertisement tools. *PLOS ONE, 18*(2), e0281243. https://doi.org/10.1371/journal.pone.0281243

Pötzschke, S., & Braun, M. (2017). Migrant sampling using Facebook advertisements: A case study of polish migrants in four European countries. *Social Science Computer Review, 35*(5), 633–653. https://doi.org/10.1177/0894439316666262

Roberts, C., Gilbert, E., Allum, N., & Eisner, L. (2019). Research synthesis: Satisficing in surveys: A systematic review of the literature. *Public Opinion Quarterly, 83*(3), 598–626. https://doi.org/10.1093/poq/nfz035

Stern, M. J., Fordyce, E., Carpenter, R., Viox, M. H., Michaels, S., Harper, C., Johns, M. M., & Dunville, R. (2022). Evaluating the data quality of a national sample of young sexual and gender minorities recruited using social media: The influence of different design formats. *Social Science Computer Review, 40*(3), 663–677. https://doi.org/10.1177/0894439320928240

Tangmanee, C., & Niruttinanon, P. (2019). Web survey's completion rates: Effects of forced responses, question display styles, and subjects' attitude. *International Journal of Research in Business and Social Science (2147-4478), 8*(1), 20–29. https://doi.org/10.20525/ijrbs.v8i1.183

Wenemark, M., Persson, A., Brage, H. N., Svensson, T., & Kristenson, M. (2011). Applying motivation theory to achieve increased response rates, respondent satisfaction and data quality. *Journal of Official Statistics, 27*(2), 393–414. https://www.proquest.com/docview/2821376428?pq-origsite=gscholar&fromopenview=true&sourcetype=Scholarly%20Journals

Zhang, C., & Conrad, F. (2014). Speeding in web surveys: The tendency to answer very fast and its association with straightlining. *Survey Research Methods, 8*(2), 127–135. https://doi.org/10.18148/srm/2014.v8i2.5453

Zillmann, D., Schmitz, A., Skopek, J., & Blossfeld, H.-P. (2014). Survey topic and unit nonresponse. *Quality & Quantity, 48*(4), 2069–2088. https://doi.org/10.1007/s11135-013-9880-y

Zindel, Z. (2023). Social media recruitment in online survey research: A systematic literature review. *methods, data, analyses, 17*(2), 207–248. https://doi.org/10.12758/mda.2022.15

Appendix

A1 Descriptive Statistics and Survey Items

Figure A1 Advertisement images used in the Facebook ads to recruit respondents for the survey with varying associations to the survey topics of immigration and climate change. None of the images used in our study were generated (or partially generated/altered) by AI.

Table A1 Campaign performance and number of started surveys in Germany and the United States

Advertisement	Impression	Reach	Unique link click	Impression to reach ratio	Cost per unique link click (in €)	Started survey	
Germany						Count	%
Immigration-strong	25,008	17,976	1,646	1.39	0.11	1275	30.6
Immigration-loose	23,371	18,728	1,116	1.25	0.16	654	15.7
Climate-strong	24,804	18,568	1,176	1.34	0.15	929	22.2
Climate-loose	28,067	22,088	1,384	1.27	0.13	578	13.9
Neutral	19,929	14,864	859	1.34	0.21	737	17.7
United States							
Immigration-strong	43,191	32,160	2,567	1.34	0.27	1345	24.6
Immigration-loose	52,410	38,592	2,684	1.36	0.26	867	15.9
Climate-strong	50,128	36,759	2,447	1.36	0.28	1428	26.1
Climate-loose	65,716	50,608	2,778	1.30	0.25	954	17.4
Neutral	27,079	18,888	1,273	1.43	0.55	875	16.0

Table A2 Data cleaning of the survey

Design	Immigration-strong	Immigration-loose	Climate-strong	Climate-loose	Neutral	Missing	Total
Germany							
Survey landing page hits	1,837	1,168	1,245	1,437	981	159	6,827
Exclusion of cases without ad information	1,837	1,168	1,245	1,437	981	0	6,668
Started surveys	1,275	654	926	578	737	0	4,170
Completed surveys	645	392	625	359	474	0	2,495
United States							
Survey landing page hits	2,757	2,819	2,547	2,763	1,520	190	12,596
Exclusion of cases without ad information	2,757	2,819	2,547	2,763	1,520	0	12,406
Started surveys	1,345	867	1,428	954	875	0	5,469
Completed surveys	559	354	716	479	412	0	2,520

Table A3 Question text measuring attitudes toward immigration in the survey for Germany and the United States used for calculation of inconsistent responses

Positive framing	
1.	Legal immigrants to America/Germany who are not citizens should have the same rights as American citizens.
2.	Immigrants are generally good for America's/Germany's economy.
3.	Legal immigrants should have equal access to public education as American citizens.
4.	Immigrants improve American/German society by bringing new ideas and cultures.
Negative framing	
1.	American/German culture is generally undermined by immigrants.
2.	Immigrants increase crime rates.
3.	America/Germany should take stronger measures to exclude illegal immigrants.
4.	Immigrants take jobs away from people who were born in America/Germany.

Table A4 Gender and age composition of the survey

(a) Gender composition

Country	Female	Male	Missing	Total
Germany	1533 (41%)	1974 (52%)	271 (7%)	3778 (100%)
United States	1881 (42%)	2208 (50%)	361 (8%)	4450 (100%)

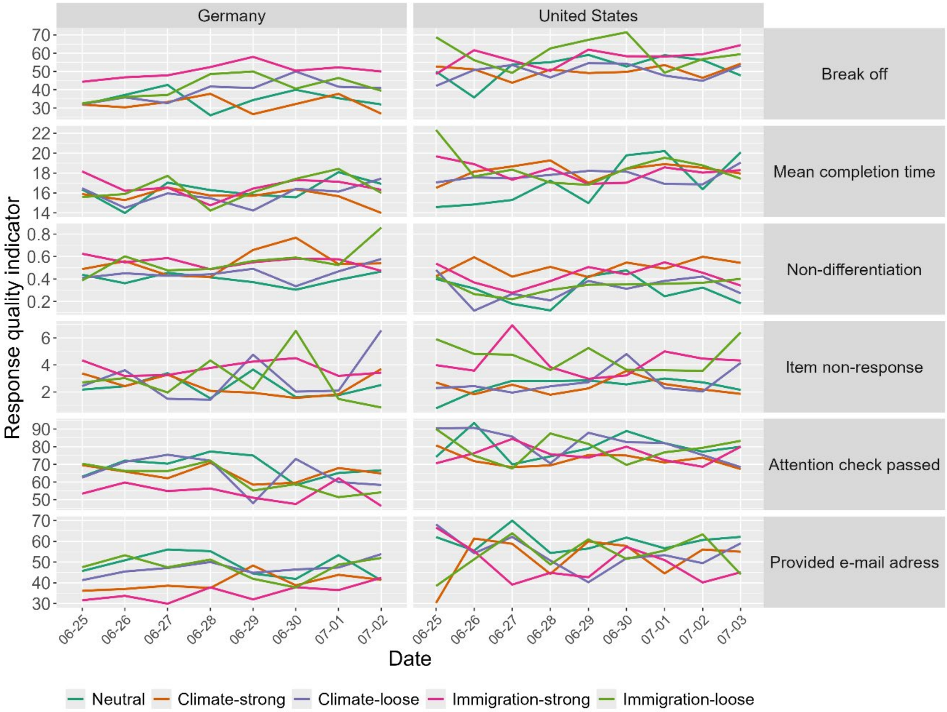
(b) Age composition

Country	*M* age	*SD*	Minimum	Maximum	Missing
Germany	60	11	18	99	10%
United States	74	10	19	96	12%

Notes: Unweighted.

Table A5 Attention check question implemented in the survey

	Questionnaire text
Question text	Here are some common statements on politics and society. Please state whether you agree or disagree.
Statements	a Politicians care about what ordinary people think. b People like me do not have any influence on the government. c Politics is so complicated people like me are not able to understand what is going on. d Please click "rather disagree." e Citizens lack possibilities to influence politics. f In a democracy it is the duty of all citizens to vote regularly in elections.
Response scale	1 strongly disagree 2 rather disagree 3 neither agree nor disagree 4 rather agree 5 strongly agree



Response quality indicator over time (daily estimates) by advertisement design and country. Unweighted.

Figure A2 Descriptive statistics for the response quality indicator over the recruitment period by advertisement design for Germany and the United States. Unweighted.

A2 Comparison of Quality Indicators with the Online Panel

In addition to the social media recruitment efforts, a commercial panel company was asked to recruit a representative sample in both countries using the same survey questions. In this baseline sample, no promotional images were employed, and no information regarding the survey topics was provided (thus corresponding to the neutral design in the social media recruitment). For the reference sample, we received 1,555 surveys for Germany and 1,576 surveys for the United States from the company that fulfilled the inclusion criteria. As expected from an online panel, almost everyone completed the survey, with only 56 respondents in Germany and 55 respondents in the United States not completing the survey (see Table A7).

A3 Regression Results

Table A6 Number of started and completed surveys from the online panel

	Germany	United States
Started surveys	1,555	1,576
Completed surveys	1,499	1,521

A quota sampling approach was employed that takes the respective population composition of Germany and the United States into account. Thus, the gender composition of the sample is balanced. The mean age of the sample is 50 years for Germany and 48 years for the United States (see Table A7). Looking at the quality indicators, we see a very low break-off rate of less than 4% in both countries. The average completion time is about 11 minutes in both countries. Non-differentiation is higher in Germany, at 0.6, than in the United States, at 0.5. The item non-response rate is 0.3 in the United States and 0.6 in Germany. The percentage of respondents passing the attention check question is quite high, at 94% in Germany and 89% in the United States (see Table A8).

Table A7 Gender and age composition of the online panel

(a) Gender composition

Country	Female	Male	Total
Germany	790 (51%)	765 (49%)	1,555 (100%)
United States	793 (50%)	783 (50%)	1,576 (100%)

(b) Age composition

Country	*M* age	*SD*
Germany	50	17
United States	48	17

Notes: Unweighted. There are no missing values as this question was mandatory.

Table A8 Descriptive statistics by country for the online panel showing the survey quality indicators: break-off rate, mean completion time, non-differentiation, item non-response, attention check

	Germany		United States	
Break-off rate (%)	3.60	[2.78, 4.65]	3.49	[2.69, 4.51]
Mean completion time (min)	11.75	[11.27, 12.23]	11.34	[10.83, 11.84]
Non-differentiation	0.63	[0.58, 0.67]	0.46	[0.41, 0.50]
Item non-response (%)	0.58	[0.54, 0.63]	0.34	[0.31, 0.38]
Attention check passed (%)	94.48	[93.21, 95.53]	88.77	[87.09, 90.26]

Notes: Values in brackets refer to the 95% confidence interval. Unweighted.

A3 Regression Results

Table A9 Stepwise regression results for Germany, outcome: break-off

Stepwise regression results for Germany, outcome: break-off					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: break-off (ref. completion)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.702 [0.432, 1.140]	0.751 [0.459, 1.227]	0.744 [0.453, 1.221]	0.747 [0.454, 1.231]	0.704 [0.426, 1.162]
Design: climate-loose	0.991 [0.589, 1.666]	1.003 [0.597, 1.687]	0.998 [0.592, 1.682]	1.037 [0.613, 1.754]	0.971 [0.573, 1.647]
Design: immigration-strong	1.665* [1.101, 2.519]	1.742** [1.148, 2.643]	1.747** [1.149, 2.656]	1.676* [1.099, 2.557]	1.584* [1.036, 2.422]
Design: immigration-loose	1.072 [0.652, 1.761]	1.131 [0.686, 1.866]	1.112 [0.674, 1.836]	1.131 [0.683, 1.874]	1.089 [0.656, 1.807]
Sex: male		0.773† [0.578, 1.034]	0.786 [0.587, 1.052]	0.877 [0.651, 1.180]	0.911 [0.676, 1.228]
Age: 40–59			1.595 [0.678, 3.752]	1.575 [0.666, 3.727]	1.585 [0.669, 3.754]
Age: 60–64			2.007 [0.834, 4.830]	1.875 [0.774, 4.539]	1.873 [0.773, 4.540]
Age: 65+			1.534 [0.645, 3.644]	1.365 [0.570, 3.267]	1.362 [0.568, 3.265]
Income: low				1.567* [1.041, 2.359]	1.617* [1.073, 2.437]
Income: medium				0.936 [0.601, 1.459]	0.939 [0.602, 1.463]
Income: missing				2.710** [1.857, 3.954]	2.766** [1.894, 4.040]
Device: mobile					2.120* [1.094, 4.109]
Constant	0.084** [0.060, 0.118]	0.093** [0.065, 0.133]	0.057** [0.023, 0.140]	0.041** [0.016, 0.105]	0.021** [0.007, 0.064]
Observations	2,455	2,455	2,455	2,455	2,455
Log likelihood	-702.957	-701.457	-699.641	-682.934	-679.937
AIC	1,415.915	1,414.914	1,417.282	1,389.868	1,385.873

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A10 Stepwise regression results for the United States, outcome: break-off

Stepwise regression results for the United States, outcome: break-off					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: break-off (ref. completion)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.992 [0.629, 1.563]	1.171 [0.732, 1.874]	1.147 [0.714, 1.843]	1.088 [0.675, 1.754]	0.982 [0.606, 1.592]
Design: climate-loose	1.017 [0.624, 1.660]	1.033 [0.633, 1.686]	0.986 [0.601, 1.616]	0.975 [0.593, 1.605]	0.878 [0.531, 1.452]
Design: immigration-strong	1.455† [0.931, 2.274]	1.518† [0.970, 2.377]	1.472† [0.938, 2.311]	1.396 [0.886, 2.201]	1.352 [0.856, 2.137]
Design: immigration-loose	1.098 [0.654, 1.846]	1.214 [0.719, 2.051]	1.168 [0.689, 1.979]	1.147 [0.674, 1.953]	0.977 [0.570, 1.674]
Sex: male		0.654** [0.488, 0.877]	0.660** [0.491, 0.886]	0.768† [0.567, 1.039]	0.833 [0.613, 1.132]
Age: 40–59			2.633 [0.327, 21.178]	2.827 [0.349, 22.878]	2.914 [0.360, 23.621]
Age: 60–64			1.951 [0.240, 15.866]	1.964 [0.240, 16.048]	2.049 [0.250, 16.782]
Age: 65+			3.043 [0.411, 22.513]	2.761 [0.371, 20.532]	2.789 [0.374, 20.778]
Income: low				0.931 [0.531, 1.631]	0.926 [0.528, 1.624]
Income: medium				0.917 [0.563, 1.492]	0.902 [0.554, 1.470]
Income: missing				2.459** [1.552, 3.895]	2.372** [1.496, 3.762]
Device: mobile					2.237** [1.392, 3.593]
Constant	0.082** [0.057, 0.118]	0.096** [0.066, 0.139]	0.034** [0.005, 0.249]	0.026** [0.003, 0.201]	0.014** [0.002, 0.110]
Observations	2,571	2,571	2,571	2,571	2,571
Log likelihood	-739.125	-735.095	-733.259	-712.284	-705.729
AIC	1,488.250	1,482.190	1,484.517	1,448.569	1,437.458

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A11 Stepwise regression results for Germany, outcome: speeding

Stepwise regression results for Germany, outcome: speeding					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: speeding (ref. not speeding)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.748 [0.500, 1.120]	0.775 [0.514, 1.168]	0.595* [0.390, 0.907]	0.592* [0.388, 0.902]	0.620* [0.405, 0.950]
Design: climate-loose	0.738 [0.459, 1.186]	0.744 [0.463, 1.196]	0.625† [0.385, 1.014]	0.618† [0.381, 1.004]	0.648† [0.398, 1.055]
Design: immigration-strong	0.766 [0.511, 1.146]	0.784 [0.522, 1.178]	0.673† [0.444, 1.021]	0.672† [0.443, 1.020]	0.700† [0.460, 1.065]
Design: immigration-loose	0.741 [0.467, 1.175]	0.764 [0.480, 1.216]	0.753 [0.468, 1.210]	0.749 [0.466, 1.205]	0.774 [0.481, 1.247]
Sex: male		0.873 [0.657, 1.161]	0.863 [0.645, 1.154]	0.840 [0.626, 1.126]	0.816 [0.607, 1.096]
Age: 40–59			0.563* [0.338, 0.937]	0.560* [0.336, 0.934]	0.558* [0.335, 0.931]
Age: 60–64			0.472** [0.270, 0.825]	0.482* [0.275, 0.844]	0.481* [0.275, 0.843]
Age: 65+			0.142** [0.077, 0.260]	0.149** [0.081, 0.274]	0.149** [0.081, 0.274]
Income: low				0.736 [0.490, 1.105]	0.718 [0.478, 1.080]
Income: medium				0.790 [0.547, 1.141]	0.788 [0.545, 1.138]
Income: missing				0.792 [0.517, 1.211]	0.778 [0.508, 1.192]
Device: mobile					0.664† [0.427, 1.033]
Constant	0.138** [0.103, 0.185]	0.146** [0.107, 0.200]	0.409** [0.232, 0.723]	0.476* [0.264, 0.861]	0.682 [0.337, 1.382]
Observations	2,247	2,247	2,247	2,247	2,247
Log likelihood	-725.305	-724.871	-692.755	-691.148	-689.606
AIC	1,460.610	1,461.742	1,403.511	1,406.296	1,405.213

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A12 Stepwise regression results for the United States, outcome: speeding

Stepwise regression results for the United States, outcome: speeding					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: speeding (ref. not speeding)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.372** [0.253, 0.548]	0.360** [0.241, 0.538]	0.424** [0.279, 0.644]	0.425** [0.279, 0.649]	0.489** [0.319, 0.751]
Design: climate-loose	0.514** [0.344, 0.766]	0.512** [0.343, 0.764]	0.712 [0.467, 1.086]	0.706 [0.462, 1.079]	0.820 [0.532, 1.262]
Design: immigration-strong	0.476** [0.322, 0.704]	0.471** [0.318, 0.698]	0.577** [0.382, 0.871]	0.578** [0.382, 0.876]	0.578* [0.380, 0.879]
Design: immigration-loose	0.420** [0.264, 0.669]	0.410** [0.256, 0.657]	0.520** [0.318, 0.850]	0.521** [0.318, 0.853]	0.677 [0.408, 1.122]
Sex: male		1.095 [0.826, 1.451]	0.978 [0.732, 1.307]	0.914 [0.680, 1.228]	0.799 [0.590, 1.081]
Age: 40–59			0.521 [0.239, 1.137]	0.501† [0.228, 1.102]	0.467† [0.210, 1.037]
Age: 60–64			0.404* [0.184, 0.887]	0.395* [0.179, 0.871]	0.352* [0.158, 0.786]
Age: 65+			0.105** [0.052, 0.215]	0.106** [0.052, 0.218]	0.098** [0.047, 0.203]
Income: low				0.617* [0.403, 0.945]	0.625* [0.407, 0.961]
Income: medium				0.613** [0.424, 0.886]	0.620* [0.427, 0.901]
Income: missing				0.418** [0.269, 0.651]	0.436** [0.279, 0.682]
Device: mobile					0.426** [0.312, 0.582]
Constant	0.219** [0.169, 0.284]	0.211** [0.159, 0.281]	1.248 [0.610, 2.555]	2.089† [0.955, 4.567]	4.078** [1.778, 9.354]
Observations	2,355	2,355	2,355	2,355	2,355
Log likelihood	-767.650	-767.450	-717.532	-709.904	-696.205
AIC	1,545.300	1,546.900	1,453.064	1,443.807	1,418.410

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A13 Stepwise regression results for Germany, outcome: non-differentiation

Stepwise regression results for Germany, outcome: non-differentiation					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: non-differentiation (ref. no non-differentiation)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	1.171 [0.906, 1.514]	1.102 [0.849, 1.432]	1.127 [0.865, 1.469]	1.135 [0.871, 1.480]	1.122 [0.859, 1.465]
Design: climate-loose	1.000 [0.741, 1.350]	0.984 [0.728, 1.329]	0.998 [0.738, 1.350]	1.010 [0.746, 1.367]	0.996 [0.735, 1.351]
Design: immigration-strong	1.285† [0.996, 1.657]	1.233 [0.953, 1.594]	1.245† [0.962, 1.612]	1.244† [0.960, 1.612]	1.231 [0.949, 1.597]
Design: immigration-loose	1.263 [0.949, 1.679]	1.206 [0.904, 1.608]	1.217 [0.912, 1.623]	1.225 [0.918, 1.635]	1.215 [0.910, 1.623]
Sex: male		1.256* [1.053, 1.497]	1.246* [1.045, 1.486]	1.289** [1.078, 1.541]	1.300** [1.086, 1.555]
Age: 40–59			0.843 [0.557, 1.278]	0.845 [0.557, 1.282]	0.846 [0.558, 1.284]
Age: 60–64			0.788 [0.509, 1.221]	0.770 [0.496, 1.196]	0.770 [0.496, 1.196]
Age: 65+			0.930 [0.610, 1.419]	0.895 [0.586, 1.369]	0.895 [0.585, 1.369]
Income: low				1.284* [1.013, 1.628]	1.294* [1.020, 1.641]
Income: medium				1.149 [0.920, 1.435]	1.150 [0.920, 1.436]
Income: missing				1.269† [0.989, 1.629]	1.274† [0.993, 1.636]
Device: mobile					1.138 [0.844, 1.534]
Constant	0.676** [0.557, 0.821]	0.612** [0.496, 0.754]	0.701 [0.449, 1.094]	0.614* [0.388, 0.974]	0.547* [0.321, 0.933]
Observations	2,201	2,201	2,201	2,201	2,201
Log likelihood	-1,505.645	-1,502.405	-1,501.147	-1,498.229	-1,497.867
AIC	3,021.289	3,016.811	3,020.293	3,020.457	3,021.734

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A14 Stepwise regression results for the United States, outcome: non-differentiation

Stepwise regression results for the United States, outcome: non-differentiation					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: non-differentiation (ref. no non-differentiation)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	1.738** [1.341, 2.253]	1.618** [1.238, 2.115]	1.608** [1.227, 2.108]	1.627** [1.240, 2.134]	1.524** [1.158, 2.006]
Design: climate-loose	1.210 [0.913, 1.603]	1.201 [0.905, 1.592]	1.175 [0.883, 1.563]	1.188 [0.893, 1.581]	1.117 [0.836, 1.491]
Design: immigration-strong	1.481** [1.130, 1.941]	1.452** [1.107, 1.905]	1.430* [1.087, 1.879]	1.440** [1.095, 1.895]	1.421* [1.080, 1.872]
Design: immigration-loose	1.318† [0.976, 1.779]	1.258 [0.929, 1.704]	1.232 [0.907, 1.673]	1.247 [0.918, 1.694]	1.133 [0.829, 1.549]
Sex: male		1.204* [1.013, 1.432]	1.208* [1.015, 1.437]	1.237* [1.036, 1.477]	1.297** [1.084, 1.553]
Age: 40–59			1.713 [0.735, 3.992]	1.758 [0.755, 4.096]	1.801 [0.771, 4.204]
Age: 60–64			1.260 [0.543, 2.925]	1.271 [0.548, 2.948]	1.322 [0.569, 3.072]
Age: 65+			1.667 [0.760, 3.657]	1.702 [0.776, 3.733]	1.726 [0.785, 3.794]
Income: low				1.252 [0.944, 1.660]	1.241 [0.936, 1.647]
Income: medium				1.035 [0.811, 1.320]	1.026 [0.803, 1.309]
Income: missing				1.117 [0.861, 1.449]	1.094 [0.842, 1.420]
Device: mobile					1.451** [1.162, 1.813]
Constant	0.500** [0.405, 0.617]	0.464** [0.372, 0.580]	0.288** [0.131, 0.634]	0.254** [0.113, 0.572]	0.190** [0.083, 0.437]
Observations	2,421	2,421	2,421	2,421	2,421
Log likelihood	-1,626.546	-1,624.329	-1,622.075	-1,620.441	-1,614.937
AIC	3,263.092	3,260.658	3,262.151	3,264.881	3,255.874

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A15 Stepwise regression results for Germany, outcome: item non-response

Stepwise regression results for Germany, outcome: item non-response					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: item non-response (ref. no item non-response)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	1.017 [0.799, 1.295]	1.088 [0.851, 1.391]	1.194 [0.929, 1.534]	1.236 [0.943, 1.622]	1.239 [0.944, 1.628]
Design: climate-loose	1.071 [0.812, 1.413]	1.085 [0.822, 1.433]	1.146 [0.865, 1.517]	1.234 [0.911, 1.670]	1.237 [0.912, 1.676]
Design: immigration-strong	1.223† [0.963, 1.553]	1.279* [1.005, 1.629]	1.356* [1.062, 1.731]	1.277† [0.979, 1.667]	1.280† [0.980, 1.672]
Design: immigration-loose	0.822 [0.627, 1.078]	0.868 [0.660, 1.140]	0.856 [0.650, 1.126]	0.853 [0.632, 1.150]	0.854 [0.633, 1.152]
Sex: male		0.773** [0.656, 0.910]	0.783** [0.664, 0.924]	0.880 [0.734, 1.054]	0.878 [0.731, 1.054]
Age: 40–59			1.960** [1.295, 2.967]	1.931** [1.231, 3.028]	1.931** [1.231, 3.028]
Age: 60–64			2.538** [1.643, 3.920]	2.455** [1.532, 3.935]	2.455** [1.532, 3.935]
Age: 65+			2.850** [1.869, 4.344]	2.685** [1.697, 4.247]	2.685** [1.698, 4.248]
Income: low				1.026 [0.818, 1.288]	1.025 [0.816, 1.287]
Income: medium				0.979 [0.789, 1.213]	0.978 [0.789, 1.213]
Income: missing				16.799** [11.481, 24.580]	16.781** [11.467, 24.558]
Device: mobile					0.972 [0.713, 1.324]
Constant	0.983 [0.819, 1.179]	1.098 [0.903, 1.335]	0.460** [0.295, 0.716]	0.309** [0.188, 0.506]	0.317** [0.180, 0.559]
Observations	2,455	2,455	2,455	2,455	2,455
Log likelihood	-1,696.657	-1,691.898	-1,674.754	-1,469.936	-1,469.920
AIC	3,403.314	3,395.795	3,367.508	2,963.873	2,965.840

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A16 Stepwise regression results for the United States, outcome: item non-response

Stepwise regression results for the United States, outcome: item non-response					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: item non-response (ref. no item non-response)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	1.233† [0.967, 1.571]	1.474** [1.144, 1.898]	1.408** [1.090, 1.819]	1.365* [1.023, 1.822]	1.393* [1.040, 1.866]
Design: climate-loose	1.210 [0.931, 1.573]	1.233 [0.947, 1.606]	1.147 [0.877, 1.499]	1.161 [0.858, 1.571]	1.182 [0.871, 1.604]
Design: immigration-strong	1.388* [1.078, 1.787]	1.459** [1.130, 1.883]	1.387* [1.072, 1.795]	1.315† [0.982, 1.761]	1.320† [0.986, 1.767]
Design: immigration-loose	1.009 [0.759, 1.341]	1.125 [0.843, 1.502]	1.061 [0.792, 1.420]	1.038 [0.746, 1.444]	1.069 [0.764, 1.496]
Sex: male		0.634** [0.537, 0.748]	0.645** [0.546, 0.762]	0.849† [0.702, 1.028]	0.837† [0.689, 1.015]
Age: 40–59			1.257 [0.557, 2.835]	1.618 [0.652, 4.015]	1.606 [0.647, 3.990]
Age: 60–64			1.512 [0.678, 3.375]	1.734 [0.705, 4.262]	1.716 [0.697, 4.226]
Age: 65+			2.118* [1.004, 4.471]	2.101† [0.907, 4.864]	2.093† [0.903, 4.854]
Income: low				1.325† [0.991, 1.773]	1.326† [0.991, 1.773]
Income: medium				1.133 [0.877, 1.462]	1.135 [0.879, 1.465]
Income: missing				12.481** [9.283, 16.781]	12.562** [9.339, 16.898]
Device: mobile					0.901 [0.717, 1.132]
Constant	0.725** [0.598, 0.880]	0.865 [0.705, 1.062]	0.451* [0.213, 0.952]	0.183** [0.077, 0.437]	0.199** [0.082, 0.483]
Observations	2,571	2,571	2,571	2,571	2,571
Log likelihood	-1,770.216	-1,755.555	-1,747.619	-1,462.862	-1,462.464
AIC	3,550.431	3,523.110	3,513.238	2,949.725	2,950.927

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A17 Stepwise regression results for Germany, outcome: attention check passed

Stepwise regression results for Germany, outcome: attention check passed					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: attention check passed (ref. attention check failed)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.826 [0.625, 1.093]	0.847 [0.637, 1.125]	0.757† [0.566, 1.012]	0.737* [0.549, 0.988]	0.775† [0.576, 1.042]
Design: climate-loose	0.881 [0.638, 1.216]	0.886 [0.642, 1.224]	0.836 [0.603, 1.159]	0.791 [0.569, 1.101]	0.841 [0.603, 1.172]
Design: immigration-strong	0.587** [0.447, 0.772]	0.597** [0.453, 0.787]	0.553** [0.418, 0.731]	0.546** [0.412, 0.725]	0.574** [0.432, 0.763]
Design: immigration-loose	0.761† [0.558, 1.038]	0.775 [0.566, 1.060]	0.777 [0.567, 1.065]	0.758† [0.551, 1.042]	0.778 [0.565, 1.072]
Sex: male		0.914 [0.757, 1.103]	0.917 [0.758, 1.108]	0.845† [0.697, 1.026]	0.817* [0.673, 0.993]
Age: 40–59			0.874 [0.558, 1.369]	0.859 [0.545, 1.352]	0.854 [0.542, 1.346]
Age: 60–64			0.518** [0.324, 0.828]	0.537* [0.334, 0.864]	0.534** [0.332, 0.859]
Age: 65+			0.524** [0.332, 0.826]	0.573* [0.362, 0.909]	0.575* [0.362, 0.912]
Income: low				0.543** [0.422, 0.699]	0.526** [0.408, 0.677]
Income: medium				0.635** [0.500, 0.808]	0.632** [0.497, 0.804]
Income: missing				0.493** [0.374, 0.650]	0.485** [0.368, 0.640]
Device: mobile					0.534** [0.376, 0.760]
Constant	2.306** [1.863, 2.854]	2.399** [1.908, 3.017]	3.800** [2.333, 6.190]	5.646** [3.380, 9.433]	9.981** [5.430, 18.347]
Observations	2,081	2,081	2,081	2,081	2,081
Log likelihood	-1,348.735	-1,348.296	-1,332.194	-1,313.826	-1,307.277
AIC	2,707.469	2,708.591	2,682.387	2,651.652	2,640.555

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A18 Stepwise regression results for the United States, outcome: attention check passed

Stepwise regression results for the United States, outcome: attention check passed					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: attention check passed (ref. attention check failed)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.559** [0.404, 0.772]	0.587** [0.421, 0.819]	0.584** [0.417, 0.818]	0.586** [0.418, 0.821]	0.666* [0.473, 0.939]
Design: climate-loose	0.888 [0.617, 1.278]	0.893 [0.620, 1.286]	0.895 [0.619, 1.295]	0.891 [0.615, 1.290]	1.004 [0.690, 1.461]
Design: immigration-strong	0.705* [0.501, 0.993]	0.718† [0.509, 1.012]	0.718† [0.507, 1.015]	0.719† [0.508, 1.018]	0.726† [0.511, 1.031]
Design: immigration-loose	0.896 [0.606, 1.323]	0.927 [0.625, 1.375]	0.927 [0.623, 1.380]	0.915 [0.614, 1.363]	1.132 [0.754, 1.698]
Sex: male		0.874 [0.701, 1.091]	0.873 [0.699, 1.089]	0.834 [0.666, 1.045]	0.745* [0.592, 0.937]
Age: 40–59			1.339 [0.543, 3.300]	1.292 [0.523, 3.193]	1.212 [0.487, 3.019]
Age: 60–64			1.421 [0.578, 3.494]	1.416 [0.575, 3.488]	1.303 [0.525, 3.237]
Age: 65+			1.195 [0.528, 2.704]	1.201 [0.529, 2.724]	1.142 [0.500, 2.611]
Income: low				0.918 [0.647, 1.303]	0.931 [0.654, 1.327]
Income: medium				1.060 [0.782, 1.436]	1.086 [0.799, 1.475]
Income: missing				0.758† [0.548, 1.049]	0.794 [0.572, 1.102]
Device: mobile					0.419** [0.310, 0.568]
Constant	4.692** [3.590, 6.133]	4.968** [3.737, 6.603]	4.096** [1.799, 9.326]	4.492** [1.900, 10.624]	9.218** [3.718, 22.852]
Observations	2,137	2,137	2,137	2,137	2,137
Log likelihood	-1,120.558	-1,119.845	-1,119.276	-1,116.045	-1,098.141
AIC	2,251.116	2,251.691	2,256.551	2,256.090	2,222.282

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A19 Stepwise regression results for Germany, outcome: provided e-mail address

Stepwise regression results for Germany, outcome: provided e-mail address					
	Odds ratio estimate for binary logistic regression, Germany				
	Outcome: provided e-mail address (ref. did not provide e-mail address)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.616** [0.483, 0.786]	0.547** [0.427, 0.703]	0.564** [0.438, 0.726]	0.552** [0.426, 0.714]	0.580** [0.447, 0.752]
Design: climate-loose	0.842 [0.638, 1.112]	0.823 [0.622, 1.088]	0.837 [0.633, 1.109]	0.810 [0.608, 1.078]	0.856 [0.642, 1.143]
Design: immigration-strong	0.504** [0.395, 0.643]	0.464** [0.363, 0.595]	0.471** [0.368, 0.604]	0.481** [0.374, 0.620]	0.503** [0.390, 0.649]
Design: immigration-loose	0.894 [0.682, 1.171]	0.815 [0.620, 1.072]	0.819 [0.622, 1.077]	0.803 [0.607, 1.062]	0.830 [0.627, 1.100]
Sex: male		1.539** [1.300, 1.821]	1.532** [1.294, 1.814]	1.436** [1.207, 1.708]	1.385** [1.163, 1.649]
Age: 40–59			0.868 [0.586, 1.285]	0.908 [0.611, 1.350]	0.905 [0.608, 1.348]
Age: 60–64			0.910 [0.602, 1.377]	0.983 [0.646, 1.495]	0.983 [0.645, 1.497]
Age: 65+			1.015 [0.681, 1.512]	1.135 [0.757, 1.701]	1.138 [0.758, 1.710]
Income: low				0.920 [0.734, 1.152]	0.891 [0.710, 1.117]
Income: medium				0.920 [0.744, 1.138]	0.916 [0.740, 1.134]
Income: missing				0.313** [0.241, 0.407]	0.304** [0.233, 0.396]
Device: mobile					0.556** [0.414, 0.746]
Constant	1.090 [0.909, 1.308]	0.907 [0.745, 1.104]	0.964 [0.633, 1.470]	1.172 [0.755, 1.818]	1.994** [1.190, 3.341]
Observations	2,455	2,455	2,455	2,455	2,455
Log likelihood	-1,661.423	-1,648.694	-1,647.303	-1,601.673	-1,593.934
AIC	3,332.845	3,309.389	3,312.606	3,227.346	3,213.869

Notes: AIC = Akaike information criterion. †$p < .10$, *$p < .05$, **$p < .01$. Unweighted. Values in brackets refer to the 95% confidence interval.

Table A20 Stepwise regression results for the United States, outcome: provided e-mail address

Stepwise regression results for the United States, outcome: provided e-mail address					
	Odds ratio estimate for binary logistic regression, United States				
	Outcome: provided e-mail address (ref. did not provide e-mail address)				
	(1)	(2)	(3)	(4)	(5)
Design: climate-strong	0.854 [0.669, 1.090]	0.790† [0.614, 1.018]	0.841 [0.651, 1.086]	0.895 [0.687, 1.166]	0.983 [0.751, 1.285]
Design: climate-loose	0.722* [0.555, 0.939]	0.716* [0.550, 0.932]	0.776† [0.594, 1.013]	0.788† [0.598, 1.038]	0.860 [0.651, 1.138]
Design: immigration-strong	0.661** [0.513, 0.853]	0.648** [0.502, 0.836]	0.689** [0.533, 0.891]	0.725* [0.556, 0.946]	0.737* [0.564, 0.964]
Design: immigration-loose	0.847 [0.637, 1.125]	0.807 [0.605, 1.076]	0.864 [0.647, 1.156]	0.890 [0.659, 1.200]	1.025 [0.755, 1.392]
Sex: male		1.221* [1.036, 1.440]	1.200* [1.017, 1.416]	1.067 [0.897, 1.271]	0.996 [0.834, 1.188]
Age: 40–59			0.399† [0.155, 1.026]	0.370* [0.141, 0.972]	0.357* [0.136, 0.940]
Age: 60–64			0.290** [0.114, 0.737]	0.273** [0.105, 0.710]	0.260** [0.100, 0.677]
Age: 65+			0.258** [0.106, 0.627]	0.276** [0.111, 0.683]	0.271** [0.109, 0.672]
Income: low				1.422* [1.074, 1.884]	1.429* [1.077, 1.895]
Income: medium				1.023 [0.807, 1.297]	1.033 [0.814, 1.311]
Income: missing				0.370** [0.287, 0.478]	0.378** [0.292, 0.488]
Device: mobile					0.583** [0.469, 0.724]
Constant	1.491** [1.227, 1.812]	1.379** [1.123, 1.693]	4.864** [1.997, 11.846]	5.880** [2.310, 14.967]	8.955** [3.456, 23.204]
Observations	2,571	2,571	2,571	2,571	2,571
Log likelihood	-1,765.883	-1,763.050	-1,754.446	-1,678.126	-1,665.927
AIC	3,541.767	3,538.100	3,526.891	3,380.252	3,357.853

Notes: AIC = Akaike information criterion. †p < .10, *p < .05, **p < .01. Unweighted. Values in brackets refer to the 95% confidence interval.

A4 R Packages Used

Package	Version	Citation
arm	1.13.1	Gelman and Su (2022)
base	4.3.2	R Core Team (2023)
ggpubr	0.6.0	Kassambara (2023)
here	1.0.1	Müller (2020)
Hmisc	5.1.1	Harrell Jr (2023)
janitor	2.2.0	Firke (2023)
kableExtra	1.4.0	Zhu (2024)
psych	2.4.1	Revelle (2024)
stargazer	5.2.3	Hlavac (2022)
tidyselect	1.2.0	Henry and Wickham (2022)
tidyverse	2.0.0	Wickham et al. (2019)

Package Citations

Firke, S. (2023). *janitor: Simple tools for examining and cleaning dirty data*. https://CRAN.R-project.org/package=janitor

Gelman, A., & Su, Y. (2022). *arm: Data analysis using regression and multilevel/hierarchical models*. https://CRAN.R-project.org/package=arm

Harrell, F. E. (2023). *Hmisc: Harrell miscellaneous*. https://CRAN.R-project.org/package=Hmisc

Henry, L., & Wickham. H. (2022). *tidyselect: Select from a set of strings*. https://CRAN.R-project.org/package=tidyselect

Hlavac, M. (2022). *stargazer: Well-formatted regression and summary statistics tables*. Social Policy Institute. https://CRAN.R-project.org/package=stargazer

Kassambara, A. (2023). *ggpubr: "ggplot2" based publication ready plots*. https://CRAN.R-project.org/package=ggpubr

Müller, K. (2020). *here: A simpler way to find your files*. https://CRAN.R-project.org/package=here

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research*. Northwestern University. https://CRAN.R-project.org/package=psych

Rodriguez-Sanchez F, & Jackson, C. (2024). _*grateful: Facilitate citation of R packages.* https://pakillo.github.io/grateful

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T. L., Miller, E., Bache, S. M., Müller, K., Ooms, J., Robinson, D., Seidel, D. P., Spinu, V. (...), Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software, 4*(43), 1686. https://doi.org

Zhu, H. (2024). *kableExtra: Construct complex table with "kable" and pipe syntax*. https://CRAN.R-project.org/package=kableExtra

Online First publication date: 2025-07-31