

Privacy Concern, Trust, and Desire for Content Personalization

Darren Stevenson^{1,3} and Josh Pasek^{1,2}

Dept. of Communication Studies¹ and Institute for Social Research²
University of Michigan

The Center for Internet and Society³
Stanford Law School

Abstract

A sizable portion of content published on websites and apps is personalized for individual users. There are both costs and benefits to personalization. Critics point out the associated costs like reductions in personal privacy linked to corresponding data collection practices or the ways in which firms algorithmically curate content to serve their own financial interests over those of users. Alternatively, given the size and speed at which digital content is produced, personalization provides a necessary filtering for an otherwise unapproachable web. It allows users to more efficiently identify the information they are seeking. Yet, it remains unclear to what extent Internet users recognize these costs and benefits? In this study we investigated how members of the general public think about personalization. Surveying a broad sample of Americans, we asked people how personalized they wanted various content (advertisements, discounts, prices, news, and social media) when encountered on websites and apps. We also assessed individuals' online privacy concerns, levels of trust in Internet firms, and Internet use to investigate the relationship between these factors and individual preference for personalization. Overall, privacy concerns do not appear to diminish support for personalization. However, trust in online firms is strongly associated with wanting personalized (vs. non-personalized) content. Additionally, heavier Internet users are substantially more likely to prefer personalization, with Internet use also moderating the influence of trust on personalization preferences. While more trusting individuals appear more favorable towards personalization, this effect was significantly impacted by Internet use. Based on these findings, we conclude that focusing on the benefits rather than the costs of personalization may be a more useful starting point in ongoing debates.

Author contact: Darren Stevenson, dstev@umich.edu, darrenstevenson.org

Online content is becoming increasingly personalized. As firms have expanded their data collection efforts and as new tools enable aggregators to link data across sites, everything from the advertisements a user sees to the top search results on Google has been enhanced to maximize personal relevance. Curation algorithms often examine items such as a user's prior browsing, geolocation, consumer file marketing data, and social network connections to determine what content should be displayed to whom. As a result, a substantial portion of what appears on websites and apps is now routinely tailored for each individual viewer.

There are both costs and benefits to personalization. Critics point out the associated problems caused by a homogeneous information environment, reductions in personal privacy brought about by corresponding data collection practices, and the ways in which firms curate content to serve their own financial interests at the expense of users. On the upside, given the size and speed at which digital content is produced, personalization provides necessary filtering for an otherwise unapproachable web, allowing users to efficiently identify the information they are seeking. Personalization thus both limits the content that users can view, in sometimes questionable ways, while providing a necessary service for those hoping to find information in online media.

It is unclear to what extent users recognize these costs and benefits. To date, widespread scholarly concern about the dangers of personalization has not elicited a widespread public backlash. Some researchers have argued that ordinary individuals are unaware of personalization's risks, citing evidence that many people do not understand the context within which data collection and curation algorithms operate (Hoofnagle & King, 2008; Hoofnagle et al., 2010). But there are reasons to think that a greater understanding may not translate into united opposition. Instead, individuals who weigh the tradeoffs may arrive at the conclusion that

personalization yields net benefits (Awad & Krishnan, 2006). And ordinary individuals do not necessarily need a perfect understanding of how the process works to make these determinations. To the extent that individuals are comfortable with data sharing practices and trust the firms that are personalizing information for them, the benefits might outweigh the risks.

In the current study, we examine how privacy concerns and trust in Internet firms relate to attitudes toward personalization. To accomplish this, we first outline some of the key costs and benefits that individuals may experience due to personalization and discuss how attitudes toward those costs and benefits might be expected to translate into support for personalized content. We then present data we collected from a broad national sample of Americans and test these expectations. Finally, we discuss what the results of these analyses imply about how individuals are thinking about personalization.

Costs of Personalization

Privacy advocates and other media critics have raised a number of concerns about personalized online content. Some have suggested that personalization will create informational echo chambers (Jamieson & Capella, 2008) or filter bubbles (Pariser, 2011). In these situations, individuals encounter a homogeneous information environment, devoid of competing viewpoints, which serves to reinforce preexisting biases (e.g., conservatives may only see conservative information). Individuals lacking diverse informational exposure thereby fail to consider important information when engaging in deliberation and making decisions (Mutz, 2006; Prior, 2007). This indictment of content personalization is rooted in the perception that a diversity of viewpoints from a variety of information sources contributes to healthier deliberation

and, ultimately, an improved democratic process (Hindman, 2008; Stroud, 2011; Kim & Pasek, 2013).

A second concern linked to content personalization surrounds the potential for an invasion of privacy (Turow, 2011), as many applications rely on potentially sensitive forms of personal information to achieve customization. In order to personalize content, media firms and others must know something about the consumer (Rossi, Schwabe, & Guimarães, 2001; Fan & Poole, 2006; Zhang, Yuan, Wang, 2014). This need for granular targeted information motivates firms to collect and store data describing unique individuals. And much of these data come from consumer file marketing companies, which often link individual records through names, addresses, and social security numbers (Tsesis, 2014).¹ Hence, to the extent that individuals are concerned about privacy, there are many reasons to be wary of personalization.

Finally, individuals might also be concerned about online personalization if they do not trust the motivations of the firms that are personalizing their content. Personalization algorithms are most commonly employed by for-profit entities – such as social media platforms, search engines, and advertisers – who are responsible to shareholders or a bottom line before their user base (Bakan, 2005). Personalization algorithms thus may not have the best interests of individuals in mind when presenting content. To the extent that commercial interests contradict the interests of Internet users, this can result in “corrupt personalization” (Sandvig, 2014), much like the corrupt segmentation of media audiences identified by Baker (2002). As a result, when individuals are wary of the motives of the companies that are personalizing their information, they would be expected to dislike the possibility. Thus, in a commercialized media system, the dangers of personalization ride on questions of trust, particularly trust in the entity responsible for curating content.

¹ Though the accuracy of these links remains an open question (see Pasek et al., 2014).

Benefits of Personalization

To date, scholars have largely ignored the rationale behind personalization in the first place when leveraging their criticisms. The scale and velocity at which the web's content is produced illustrate the basic need for information filtering. And the need for a gatekeeping process is far from novel; traditional media systems relied on journalistic practices and human editors to pare down the cacophony of each day's happenings into the daily news (Lewin, 1947; White, 1961, McCombs et al., 1976). That curation must and should be done is largely uncontested (cf. Pariser, 2011). The question then becomes how curation should be achieved now that content is being produced on a scale too-large for the traditional editorial process. And similar processes apply to advertisements, which need to narrowly focus their messages in order to compete. Here too, no human could parse the wealth of data that marketers currently consider.

Even beyond the filtering necessary to extract signal from an overload of information, individuals can derive considerable benefits from well-targeted content. As personalization algorithms generate, filter, or otherwise determine what is presented online, this selective content is more likely to be of relevance for each individual Internet user (Liu et al., 2012). Indeed, sites like Facebook would be unwieldy if every user needed to sort through every post from every friend (Liu, Belkin, & Cole, 2012; Bernstein et al. 2013; Eslami et al., 2015). And though individuals may find it odd when they are followed from site to site with the sneakers they considered on Zappos, these shoes are often flanked by similar pairs that make comparison shopping easier, avoiding a paradox of choice (Schwartz, 2004). For frequent web users, the relevance conferred by personalized information can make browsing a far more efficient process.

The Importance of Personalization Preferences

Although personalization has become a topic of increasing interest and debate, little research has examined how the public thinks about this process. Among those who have taken up this topic, more common are broad conceptual debates regarding the ethics and impacts of algorithmic personalization (Bozdag & Timmermans, 2001; Bozdag, 2013; Ashman et al., 2014; Sandvig et al., 2015; Striphas, 2015). Studies examining users' preferences for information disclosure (Culnan, 1993; Phelps, Nowak, & Ferrell, 2000; Lederer, Mankoff, & Dey, 2003) and personal privacy controls (Sheehan & Hoy, 2000; Cvreck, 2006) are well established. Less attention has been given to personalization that results both from these disclosures and various privacy configurations intended to limit how personal data is collected and used. Previous work fails to offer substantive explanations for how Internet users think about personalization. Thus, despite the growing prevalence of personalization, we currently know little about the underlying mechanisms contributing to attitudes towards personalization.

There are reasons, however, to think that the way that ordinary individuals view personalization will shape how companies use these services. As firms seek to understand and respond to demands, expressed preferences play an important role in the design and features of emerging media systems. Many efforts seek to incorporate user preferences into the design process directly (e.g., Gulliksen et al., 2003; Garrett, 2010),² requiring in-depth understanding of how individuals think about the tools and systems they use.

Hence, this study attempts to unpack *who* desires personalization and *why* these individuals prefer a customized web? To do so, we concentrate on the user's point of view and examine the impacts of privacy concern, trust, and Internet use to explain why some people desire personalization and others do not. While privacy concerns are conceptually related to

² Often referred to as *user-centered design*.

trust, the two are not the same with each conjuring different underlying constructs (Metzger, 2004; Liu, 2005). This distinction is important theoretically for understanding how individuals think about personalization and therefore we treat them separately in our investigation. We incorporate the impacts of Internet use as the salience of personalization must be accounted for in assessing how individuals think about this phenomenon.

Few have studied the underlying factors that influence why individuals prefer personalization. In a related study, Chellappa & Sin (2005) make explicit the connection between privacy, trust, and online personalization. From an in-person survey (N=243) of e-commerce consumers, the authors located a negative association between privacy concern and intent to use e-commerce (and other services that rely on personalization) and a positive association between trust and the likelihood of individuals to use personalization systems. While we focus on similar variables in addition to other, our theoretical motivations and study differ in two notable ways: operationalization and context. First, Chellepa & Sin rely on an incongruous two-item scale to assess their outcome variable—likelihood of using personalization services. Specifically, researchers asked consumers about their: 1) comfort in providing personal information to firms for use in personalization and 2) comfort in using e-commerce. The former includes two constructs (information disclosure and personalization); the latter equates use of e-commerce with personalization. The authors' measure is then used as a *proxy* for likelihood to use personalization. Alternatively, we aimed to directly assess personalization's psychological antecedents in the form of expressed preferences for personalization, asking respondents "How personalized..." they would like various types of online content. This is substantively and theoretically different from relying on comfort in e-commerce transactions as a proxy for attitude towards personalization. The second main difference in our studies regards context, with the

2005 study taking place a decade ago at a time when personalization was arguably less salient for general Internet users than it is today. While we aim to build on this work, we see these differences in construct operationalization and context as theoretically important to understanding attitudes towards online personalization.

The Current Study

Given that personalization has both costs and benefits, individuals likely vary in their desire for personalization as a function of the relative salience of these tradeoffs. Two of the central costs of personalization – concern about privacy and firm trust – would be expected to vary across individuals in ways that might influence their informational preferences.³ Similarly, the benefits can be derived from personalization are likely to be contingent on the extent to which individuals engage with personalize-able media (i.e. use the Internet). Thus there are good reasons to predict individuals should vary in how personalized they think various types of content should be.

Specifically, individuals who are concerned about privacy risks should be particularly wary of personalized content. As personalization is intrinsically linked to knowing information about users, those with more aversion to personal information disclosure should be more averse to personalization if and when they connect these activities. These individuals approach control over their personal privacy as a right rather than a privilege (Goodwin, 1991). Prior investigations have made this privacy-personalization relationship explicit seeking to understand how the two are related and concluding that privacy has a negative relationship with likelihood to use tools and services that leverage personalization (Panjwani, 2013).

³ In contrast, concerns about media diversity are principally societal in nature and would not necessarily correlate with individuals' preferences or behaviors (cf. van Cuilenburg, 2007).

H1: Privacy concern will predict decreased desire for online content personalization.

Trust and specifically trust in Internet firms should also influence preferences for personalization but in a positive direction. The facilitating role of trust in adoption of online products and services is well studied phenomenon. The majority of work indicates a positive association between trust and willingness to use digital platforms (Yoon, 2002; Beldad, de Jong, & Steehouder, 2010; Kim & Kim, 2005; Weisberg, Te'eni, & Arman, 2011; Schneier, 2012). We have no reason to think the role of trust would be substantively diminished in the case of user preference for online content personalization platforms. If anything, due to the potentially sensitive nature of personal data used we anticipate this relationship to be even stronger in the case of personalization. Conversely, distrust should have the opposite affect. Internet users who are more distrusting of companies should be more concerned with the costs of personalization and consequently less likely to perceive and value its benefits. In this way, we anticipate trust to contribute to higher esteem of personalization's benefits just as distrust leads individuals to be more likely focus on negative costs.

H2: Trust in commercial websites and apps will predict increased desire for online content personalization.

Additionally, individuals who use the Internet heavily should be the most supportive of personalization. First and foremost, individuals who use the Internet more simply have more to gain from the efficiencies brought about through content personalization. Conversely, for those

who the Internet holds a smaller place in their lives the benefits of personalization and its efficiencies are of less consequence. The second rationale for why Internet use should associate with desire for personalization hinges on fundamentals of new technology adoption. For instance, in prior work the popular Technology Acceptance Model (TAM) (Davis, 1989) and its various iterations (e.g., Wu & Wang, 2005) have been used to illustrate how factors like perceived usefulness and ease of use influence both the initial adoption and frequency of use for particular technologies, including the Internet (Lederer, 2000). Similarly, in the case of acceptance and subsequent preference for personalization technologies, individuals who already widely accept and use the Internet at greater frequency should exhibit less barriers to adoption of personalization given the overlap and interdependency between these inseparable technologies.

H3: Internet use will predict increased desire for online content personalization.

There are good reasons to think that the influences of both privacy concerns and trust will be moderated by Internet use. First, those who trust Internet firms a lot are likely to be more supportive of the benefit of personalization but only if the Internet already plays a large role in their daily lives. Similarly, those who are more distrusting of companies are likely to be more focused on the costs of personalization but again only if the Internet holds prominence in their life. Second, the same should be the case for individuals more concerned about their online privacy who might focus on the downsides of personalization if the using the Internet is a large part of their lives. Similarly, high privacy conscious individuals who are not heavy Internet users have less to lose from the reduction in privacy attributed to personalization and therefore should be less likely to be concerned about personalization. For each of these potentially moderating

factors, as the risks or benefits become more prominent to users people are likely to interpret them as more problematic or advantageous, respectively. This follows a range of work illustrating that assessments of risk/reward change as these risks/rewards increase in saliency (Ajzen & Fishbein, 1980; Petty & Krosnick, 1995; Pligt & De Vries, 1998).

H4a: The influence of concerns about online privacy on desire for personalization should be stronger among heavy Internet users.

H4b: The influence of trust on desire for personalization should be stronger amount heavy Internet users.

Data

Survey data was collected from a five wave panel using a broad national sample of U.S. Americans. Panel respondents were recruited via email invitation and/or solicitation on the Clear Voice Surveys dashboard. Following participant recruitment survey data was collected by Qualtrics. In line with predetermined respondent quotas and anticipated attrition, the five wave panel contained 3,730, 2,455, 2,268, 1,047, and 819 respondents, respectively, in the individual waves. Survey waves were evenly spaced (roughly) between September-December 2014. Our study described here corresponds to responses from 736 respondents who fully completed the final wave and uses measures included in both wave 1 (demographics) and wave 5 (variables of interest).

Procedures

Anonymous participants were presented with a series of questions on a computer screen in a web browser. Participants completed the survey on their own time and location of choice while using their personal computing devices. All participants were presented with the same survey questions.

We used ordinary least squares regression to assess how individual preferences for online content personalization were related to demographics (age, gender, race, education, income), internet use, and two measures specific to personalization: trust in commercial websites/apps and online privacy concern. Due to missingness in some demographic measures for some individuals, all predictors were imputed using multiple imputation. Five datasets were produced using multiple imputation via chained equations (mice) and the results of separate analyses were pooled to produce the estimates shown.

Measures

Desire for online content personalization

We assessed the degree to which individuals say they prefer online content on websites and apps that has been personalized for them compared to non-personalized content (Cronbach's $\alpha = .90$). Despite content personalization being quite common, the degree to which Internet users are familiar with personalization both conceptually and technically varies. Accordingly, prior to responding to items for this measure participants were presented with a brief explanation of what online content personalization is and how it functions conceptually. Prior to responding to questions asking about desire for personalization, respondents were presented with the following prompt: Some websites and apps personalize the content you see using information about you ...

Information used to personalize what you see includes, but is not limited to, things like: websites you've visited or apps you've used; your age, income, marital status, race/ethnicity, political affiliation, or location; purchases you've made online or in a store; which device or software you are using to access the Internet. Respondents were then asked the following: Indicate how personalized you would like the following items when you see them on websites or apps. ... political advertisements, news stories, friends' social media posts prices, discounts, advertisements for products and services (Not at all personalized, A little personalized, Somewhat personalized, Very personalized, Extremely personalized).

Online Privacy Concern

To gauge individuals' privacy concerns specific to Internet use, we used a shortened version of Karson's (2002) online privacy concern scale ($\alpha = .87$). To provide a more robust measure, we used item-specific response options rather than the original Likert scales. Respondents were asked the following: How concerned are you about websites or apps collecting information about you? (Not at all concerned, A little concerned, Somewhat concerned, Very concerned, Extremely concerned); How important is it to you that websites or apps spare no expense to secure their computer databases that store information about you? (Not at all important, A little important, Somewhat important, Very important, Extremely important); How important is it to you that websites or apps double-check the accuracy of the information about you that they store? (Not at all important, A little important, Somewhat important, Very important, Extremely important); How important is it to you that websites or apps have procedures to correct errors in the information about you they collect? (Not at all important, A little important, Somewhat important, Very important, Extremely important); How important is it

to you that websites or apps do NOT sell information about you to other companies? (Not at all important, A little important, Somewhat important, Very important, Extremely important); How concerned are you about websites or apps using information about you for unauthorized purposes? (Not at all concerned, A little concerned, Somewhat concerned, Very concerned, Extremely concerned); How much does it bother you when a website or app collects information about you? (Does not bother me at all, Bothers me a little, Bothers me a moderate amount, Bothers me a lot, Bothers me a great deal); How much should websites or apps increase what they already do to ensure information about you stored in their computer databases is not accessed by unauthorized people? (Do not increase what they already do, Increase what they already do a little, Increase what they already do a moderate amount, Increase what they already do a lot, Increase what they already do a great deal).

Trust in Online Firms

We use Bhattacharjee's (2002) scale measuring trust in online firms. We shortened and adapted this scale to match our constructs. More specifically, Bhattacharjee validated his scale using questions that asked about specific online firms (e.g. Amazon). Rather than asking about specific firms, we measured trust in online firms more generally, prompting respondents to consider "For the commercial websites and apps that you use, ..." for a number of items related to trust ($\alpha = .93$). To provide a more robust measure, we used item-specific response options instead of the original Likert scale. Respondents were asked the following: For the commercial websites and apps that you use, how fair are they in the way they use information about you? (Not at all fair, A little fair, Somewhat fair, Very fair, Extremely fair); For the commercial websites and apps that you use, how fair are their Terms of Service (ToS) agreements? (Not at all

fair, A little fair, Somewhat fair, Very fair, Extremely fair); For the commercial websites and apps that you use, how fair are they in the way they interact with you? (Not at all fair, A little fair, Somewhat fair, Very fair, Extremely fair); For the commercial websites and apps that you use, how trustworthy are they? (Not at all trustworthy, A little trustworthy, Somewhat trustworthy, Very trustworthy, Extremely trustworthy); For the commercial websites and apps that you use, how often do they act in your best interests? (Never act in my best interests, Rarely act in my best interests, Sometimes act in my best interests, Often act in my best interests, Always act in my best interests); For the commercial websites and apps that you use, how often do they try to address your concerns? (Never try to address my concerns, Rarely try to address my concerns, Sometimes try to address my concerns, Often try to address my concerns, Always try to address my concerns); For the commercial websites and apps that you use, to what extent are they receptive to your wishes? (Not at all receptive to my wishes, A little receptive to my wishes, Somewhat receptive to my wishes, Very receptive to my wishes, Extremely receptive to my wishes)

Internet Use

Investigating online participation and web-based mobilization, Vissers et al. (2012) developed a scale to assess individuals' Internet skills by asking how often they performed a number of online activities and then simply using these reported frequencies as a proxy for skill. Differently, as these questions ask about frequency, we implemented this scale more directly to assess Internet use (rather than skill) ($\alpha = .82$). We performed minimal updates to questions to better reflect the current context. For instance, we changed the previous “[Post] messages in chat rooms, newsgroups, blogs, or in online forums” to the more contemporary “Post a message on a

blog, social media site, or other online forum.” Individual questions included: How often do you do the following activities? ... Use a search engine to find information; Make a webpage or create a blog; Use peer-to-peer file sharing to exchange music, movies, documents, etc.; Use a website or app for banking; Play an interactive game on a website or app; Post a message on a blog, social media site, or other online forum; Send an e-mail; Use a website or app to pay bills; Make an online purchase using a website or app; Make a voice or video call via the Internet (e.g. Skype, FaceTime, Google Hangouts, etc.) Response options for all items include: Never, Less than once a month, 1-3 times per month, About once a week, 2-3 times per week, Most days, Multiple times per day.

Demographics

In addition to these substantive variables, we also controlled for five demographic questions in all analyses. These included age, gender, race, education, and income. Full question wordings and distributions for these measures are shown in the Appendix (forthcoming).

Results

Distributions

Individuals in our sample were not particularly enthusiastic about personalization. Respondents indicated that they wanted no personalization at all in 39.7% of responses to personalization questions and that they wanted things to be “extremely” personalized for only 6.2% of responses. They were the most enthusiastic about personalized discounts, which 27.8% of individuals wanted either “very” or “extremely” personalized and were least enthusiastic about personalized political advertisements, which 56.2% of respondents did not want at all.

Overall, the average respondent wanted things between “a little” and “somewhat” personalized, scoring a .32 on our composite index.

In line with this seeming disinterest in personalization, respondents were generally concerned about their online privacy. Around half of respondents reported that it was “extremely important” for companies to do more to secure their data (51.9%) and to not sell their data (49.3%). The average respondent scored a .70 on this measure.

Interestingly, however, despite widespread concerns about online privacy, respondents reported they were moderately trusting of online firms, averaging a .46 across the seven items. And the extent to which individuals reported trusting online firms was unrelated to their privacy concerns (Pearson’s $r = .03, p = .43$).

The individuals in our sample tended to regularly engage in only a handful of the Internet use activities that we measured ($M = .32$). Internet use, however, was moderately correlated with trust in online firms ($r = .37, p < .001$) and slightly positively correlated with concerns about online privacy ($r = .10, p = .008$).

Predicting Preferences for Personalization

We did not find evidence for the expectation that privacy concerns would diminish support for personalization (H1). In contrast to this hypothesis, individuals who reported that they were very concerned about privacy were indistinguishable from unconcerned individuals in their desires for personalization ($b = -.004, SE = .04, p = .91$; Table 1, column 1, row 1). And this was not simply a function of multicollinearity, as there was also no zero-order correlation between these items ($r = .01, p = .69$). This suggests that privacy is not a strong deterrent for personalization desires.

The hypothesized relation between trust in online firms and desire for personalization was present, however (H2). Compared to the least trusting individuals, those who reported the greatest trust in online firms were much more likely to desire personalized content ($b = .42$, $SE = .04$, $p < .001$; Table 1, column 1, row 2). Hence, it seems that people are much more likely to decide whether they want personalization based on their willingness to trust particular services than based on more general privacy concerns.

We also found evidence that the heaviest Internet users were the most likely to desire personalization; this fell in line with the expectation that these individuals would experience the greatest gains in efficiency from well-targeted information (H3). Indeed, those who used the Internet most were the most likely to say that they wanted to see additional personalization ($b = .40$, $SE = .06$, $p < .001$; Table 1, column 1, row 2).

The potential for efficiency gains as a function of use led us to expect that Internet use might moderate the influence of privacy concerns and trust in online firms (H4). Further emphasizing our failure to confirm H1, online privacy concerns remained unrelated to personalization desires both on their own and in interaction with Internet use when this additional term was included (H4a; Table 1, column 2).

Internet use did seem to moderate the influence of trust on personalization desires (H4b). Figure 1 shows how personalization would be expected to vary across trust and Internet use for a typical individual when holding all covariates constant. Although the most trusting individuals were always more favorable toward personalization, this was clearly extenuated by Internet use, and vice-versa.

Table 1 - OLS Regressions Predicting Desire for Personalization

| | Main Effect Model | Interaction Model |
|---------------------------------|-------------------|-------------------|
| | <i>b (SE)</i> | <i>b (SE)</i> |
| Online Privacy Concerns | .00 (.04) | .08 (.08) |
| Trust in Online Firms | .42 (.04)*** | .29 (.08)*** |
| Internet Use | .40 (.06)*** | .36 (.17)* |
| Internet Use x Privacy Concerns | -- | -.29 (.24) |
| Internet Use x Trust | -- | .44 (.20)* |
| Female | -.02 (.02) | -.02 (.02) |
| Age | -.11 (.05)* | -.11 (.05)* |
| Black, Non-Hispanic | .06 (.04) | .05 (.04) |
| Hispanic | .01 (.04) | .01 (.04) |
| Other/Multiple, Non-Hispanic | .03 (.04) | .03 (.04) |
| H.S. Degree | -.03 (.07) | -.03 (.07) |
| Some College | -.04 (.07) | -.04 (.07) |
| College Degree | -.07 (.08) | -.06 (.08) |
| Graduate Degree | -.06 (.08) | -.06 (.08) |
| Income | .01 (.04) | .01 (.04) |
| Intercept | .11 (.08) | .12 (.09) |
| N | 736 | 736 |
| R-squared | .28 | .29 |

Standard errors shown in parenthesis. * $p < .05$, ** $p < .01$, *** $p < .001$

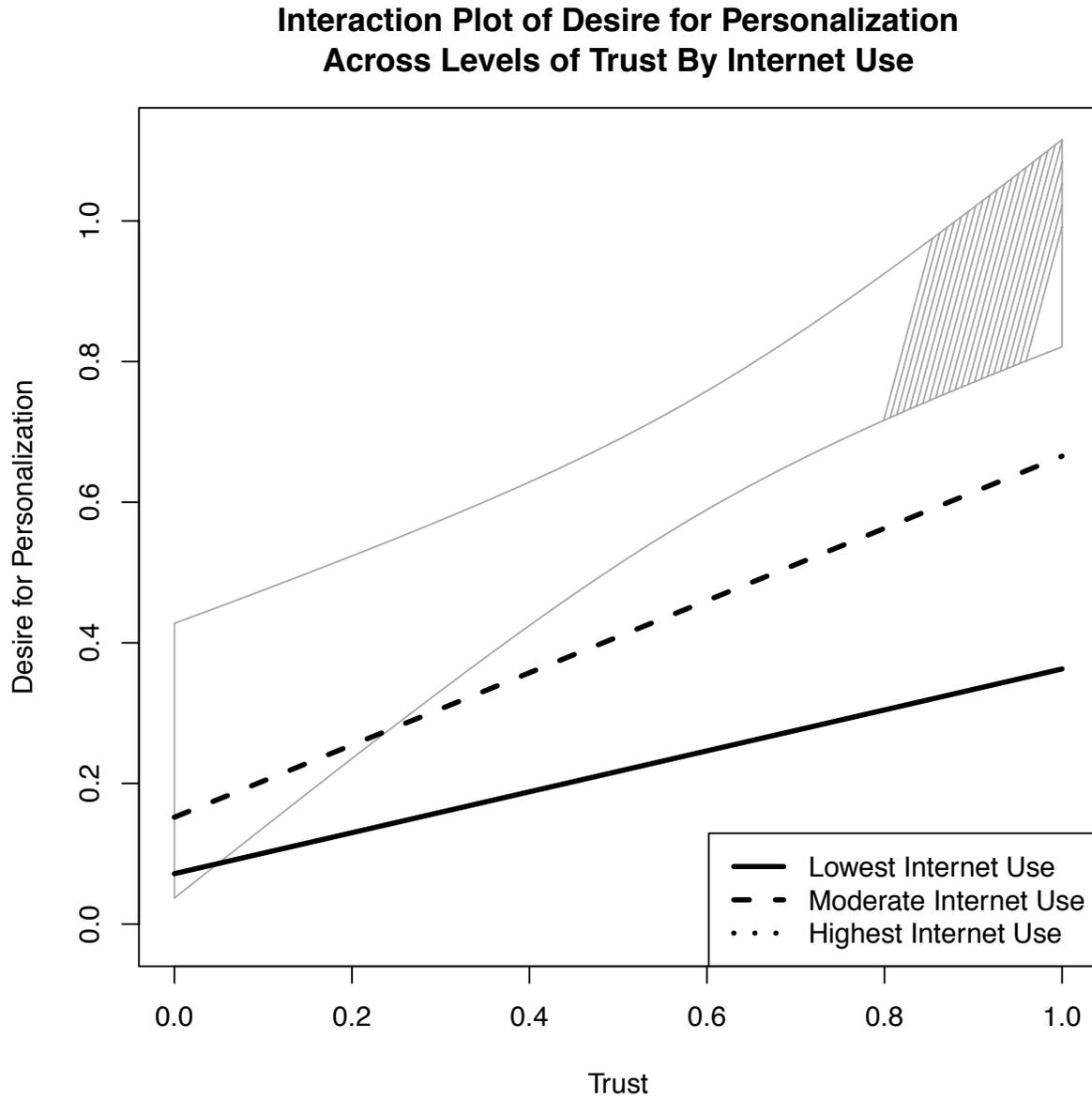


Figure 1. Internet use appears to moderate influence of trust on personalization preference (H4b).

Discussion and Conclusions

A substantial portion of content appearing on websites and apps is now personalized for each individual viewer. Therefore, there is no singular “the Web” but rather countless personalized webs each offering a unique user experience. The current study is among the first to examine how members of the public think about this phenomenon. To do so, we investigated whether Internet users say they want personalization and, if so, if this preference was reflected across the board. We assessed individuals’ privacy concerns, levels of trust in commercial firms, and Internet use to investigate the relationship between these factors and stated preferences for different types of personalized online content, including advertisements, discounts, prices, news, and social media content.

Despite considerable scholarly concern over the data collection and aggregation practices that enable personalized content, we find little evidence that privacy worries motivate personalization preferences. Instead, our evidence suggests that individuals’ preferences for personalized content are more closely rooted in the benefits of personalization than in its risks. Individuals tend to desire personalization when it seems personally beneficial and to eschew it when the benefits are unclear. And the benefits are clearest when individuals use Internet services heavily and trust the sites that provide them with content.

Our results are constrained by several factors. Notably, while we draw from a demographically-diverse sample of Americans our respondents are not statistically representative of the population and results cannot be interpreted as such. Further, respondents self-selected to participate based on a nominal financial incentive. Conceptually speaking, stated preferences for

personalization may suffer from social desirability,⁴ a common problem and one that has been observed previously in attempting to measure online privacy preferences (Berendt, Günther, & Spiekerman, 2005). Finally, we assessed *how personalized* individuals wanted various types of online media. However, this may not be how people think about personalization; that is, along a spectrum. Rather, individuals may consider their tastes for personalization in a more binary fashion—either personalized or not personalized. In our capturing the strength of this preference, we may be straying from individuals' preconceived preferences. Another potential limitation might be that respondents still misunderstand what online content personalization is and how it is achieved, despite efforts to briefly explain the process.

While the costs of personalization represent established concerns in the literature, they also correspond to somewhat idyllic conceptions of an online media environment, one where Internet users: 1) remain largely anonymous, 2) have their best interests known and upheld by firms, even when these interests oppose financial incentives of the firm providing a product or service, and 3) find themselves exposed to a perfect diversity and balance of information, opinions, and biases. In practice this ideal amounts to a paradox. Online anonymity typically opposes both having one's interests known and upheld as well as tracking an individual's media diet to ensure a proper balance of diversity. Furthermore, scholars have taken each of these ideals for granted both as normatively positive and also simply as what Internet users want.

Our results indicate that focusing on the costs may be a poor approach for scholarly inquiry as well as attempts to engage the public on this issue. Similarly, firms seeking to leverage personalization may also benefit from engaging users with a more complete set of tradeoffs highlighting the benefits which it appears individuals do consider in forming preferences for

⁴ Although in which direction we hesitate to say, as it is unclear if desiring personalization is normatively good or bad?

personalization. The problem with focusing on costs may be explained by the relatively low awareness by ordinary consumers regarding how personalization functions. That is, if individuals are largely unaware of the degree to which information about them is collected and used to customize the content they see online, then costs such as reductions in privacy have little impact on attitude formation. For instance, people may not directly associate personalization with privacy concerns and, consequently, those individuals who are more sensitive to online privacy issues may still report a desire for content personalization as indicated in our survey.

Overall, our study attempts to advance the conversation surrounding online content personalization and offer insights into why some individuals prefer personalization over others. By focusing on the user, we located the importance of considering not only the costs of personalization—which appear to have little impact on how individuals are thinking personalization—but also to consider the attitudinal impacts of personalization’s benefits for understanding how individuals form preferences in this area.

Acknowledgements

Thanks to Christian Sandvig and multiple anonymous reviewers for helpful comments and suggestions. This project was supported in part by a Winthrop B. Chamberlain Scholarship for Graduate Student Research, Department of Communication Studies, University of Michigan.

References

- Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.
- Ashman, H., Brailsford, T., Cristea, A. I., Sheng, Q. Z., Stewart, C., Toms, E. G., & Wade, V. (2014). The ethical and social implications of personalization technologies for e-learning. *Information & Management*, 51(6), 819-832.
- Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: an empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS quarterly*, 13-28.
- Beldad, A., De Jong, M., & Steehouder, M. (2010). How shall I trust the faceless and the intangible? A literature review on the antecedents of online trust. *Computers in Human Behavior*, 26(5), 857-869.
- Berendt, B., Günther, O., & Spiekerman, S., (2005). "Privacy in E-Commerce: Stated Preferences Vs. Actual Behavior," *Communications of the ACM*, 48 (4), 101–106.
- Bernstein, M. S., Bakshy, E., Burke, M., & Karrer, B. (2013, April). Quantifying the invisible audience in social networks. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 21-30). ACM.
- Bhattacharjee, A. (2002). Individual trust in online firms: Scale development and initial test. *Journal of management information systems*, 19(1), 211-241.
- Bozdag, E. (2013). Bias in algorithmic filtering and personalization. *Ethics and Information Technology*, 15(3), 209-227.

- Bozdag, V. E., & Timmermans, J. F. C. (2001, September). Values in the filter bubble Ethics of Personalization Algorithms in Cloud Computing. In 1st International Workshop on Values in Design—Building Bridges between RE, HCI and Ethics, Lisbon, Portugal.
- Chellappa, R. K., & Sin, R. G. (2005). Personalization versus privacy: An empirical examination of the online consumer's dilemma. *Information Technology and Management*, 6(2-3), 181-202.
- Culnan, Mary (1993), "How Did They Get My Name?" An Exploratory Investigation of Consumer Attitudes Toward Secondary Information Use," *MIS Quarterly*, 17 (3), 341–363.
- Cvrcek, D., Kumpost, M., Matyas, V., & Danezis, G. (2006, October). A study on the value of location privacy. In *Proceedings of the 5th ACM workshop on Privacy in Electronic Society* (pp. 109-118). ACM.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology, *MIS Quarterly* 13(3): 319–340.
- eMarketer (2014). U.S. Programmatic Ad Spend Tops \$10 Billion This Year, to Double by 2016. Available at <http://www.emarketer.com/Article/US-Programmatic-Ad-Spend-Tops-10-Billion-This-Year-Double-by-2016/1011312>
- Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., Hamilton, K., & Sandvig, C. (2015, April). I always assumed that I wasn't really that close to [her]": Reasoning about invisible algorithms in the news feed. In *Proceedings of the 33rd Annual SIGCHI Conference on Human Factors in Computing Systems* (pp. 153-162).

- Fan, H., & Poole, M. S. (2006). What is personalization? Perspectives on the design and implementation of personalization in information systems. *Journal of Organizational Computing and Electronic Commerce*, 16(3-4), 179-202.
- Garrett, J. J. (2010). *The elements of user experience: User-centered design for the web and beyond*. Berkeley, CA: Pearson Education.
- Goodwin, C. (1991). Privacy: Recognition of a consumer right. *Journal of Public Policy & Marketing*, 149-166.
- Gulliksen, J., Göransson, B., Boivie, I., Blomkvist, S., Persson, J., & Cajander, Å. (2003). Key principles for user-centred systems design. *Behaviour and Information Technology*, 22(6), 397-409.
- Hindman, M. (2008). *The Myth of Digital Democracy*. Princeton, NJ: Princeton University Press.
- Hoofnagle, C. J., King, J., Li, S., & Turow, J. (2010). How different are young adults from older adults when it comes to information privacy attitudes and policies?. Available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1589864
- Hoofnagle, C. J., & King, J. (2008). What Californians understand about privacy online. Available at <http://ssrn.com/abstract=1262130>
- Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press.
- Kim, D. H., & Pasek, J. (2013). Value-Trait Consistency in News Media Exposure. Paper presented at the 38th Annual Conference of the Midwest Association for Public Opinion Research (MAPOR), Chicago, IL.

- Kim, Y. H., & Kim, D. J. (2005, January). A study of online transaction self-efficacy, consumer trust, and uncertainty reduction in electronic commerce transaction. In Proc. of the 38th Annual Hawaii International Conference on System Sciences (HICSS). IEEE.
- Lederer, S., Mankoff, J., & Dey, A. K. (2003, April). Who wants to know what when? privacy preference determinants in ubiquitous computing. In CHI'03 extended abstracts on Human factors in computing systems (pp. 724-725). ACM.
- Lederer, A. L., Maupin, D. J., Sena, M. P., & Zhuang, Y. (2000). The technology acceptance model and the World Wide Web. *Decision support systems*, 29(3), 269-282.
- Lee, D. J., Ahn, J. H., & Bang, Y. (2011). Managing consumer privacy concerns in personalization: A strategic analysis of privacy protection. *MIS Quarterly*, 35(2), 423-444.
- Lewin, K. (1947). Frontiers in group dynamics II. Channels of group life; social planning and action research. *Human relations*, 1(2), 143-153.
- Lin, J., Amini, S., Hong, J. I., Sadeh, N., Lindqvist, J., & Zhang, J. (2012, September). Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing (pp. 501-510). ACM.
- Liu, C., Marchewka, J. T., Lu, J., & Yu, C. S. (2005). Beyond concern—a privacy-trust-behavioral intention model of electronic commerce. *Information & Management*, 42(2), 289-304.
- Liu, C., Belkin, N. J., & Cole, M. J. (2012, August). Personalization of search results using interaction behaviors in search sessions. In Proceedings of the 35th international ACM

- SIGIR conference on Research and development in information retrieval (pp. 205-214).
ACM.
- McCombs, M. E., & Shaw, D. L. (1976). Structuring the “unseen environment”. *Journal of Communication*, 26(2), 18-22.
- Metzger, M. J. (2004). Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer-Mediated Communication*, 9(4).
- Mutz, D. (2006). *Hearing the Other Side: Deliberative Versus Participatory Democracy*. New York: Cambridge University Press.
- Osorio, C., & Papagiannidis, S. (2014). Main Factors for Joining New Social Networking Sites. In F. Nah (Ed.), *HCI in Business, Lecture Notes in Computer Science*, Vol. 8527 (pp. 221-232). Switzerland: Springer International Publishing.
- Panjwani, S., Shrivastava, N., Shukla, S., & Jaiswal, S. (2013, April). Understanding the privacy-personalization dilemma for web search: A user perspective. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 3427-3430). ACM.
- Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.
- Pasek, J., Jang, S. M., Cobb, C. L., Dennis, J. M., & Disogra, C. (2014). Can Marketing Data Aid Survey Research? Examining Accuracy and Completeness in Consumer-File Data. *Public Opinion Quarterly*, 78(4), 889-916.
- Petty, R.E. & Krosnick, J.A. (1995). *Attitude strength: Antecedents and consequences*. Mahwah, NJ, Lawrence Erlbaum.
- Phelps, J., Nowak, G., & Ferrell, E., (2000). “Privacy Concerns and Consumer Willingness to Provide Information,” *Journal of Public Policy and Marketing*, 19 (1).

- Pligt, J. V. D., & De Vries, N. K. (1998). Expectancy-value models of health behaviour: The role of salience and anticipated affect. *Psychology and Health*, 13(2), 289-305.
- Prior, M. (2007). *Post-Broadcast Democracy: How Media Choice Increases Inequality in Political Involvement and Polarizes Elections*. New York: Cambridge University Press.
- Rossi, G., Schwabe, D., & Guimarães, R. (2001, April). Designing personalized web applications. In *Proceedings of the 10th international conference on World Wide Web* (pp. 275-284). ACM.
- Sandvig, C., Hamilton, K., Karahalios, K., & Langbort, C. (2015). Can an Algorithm be Unethical? Paper presented to the 65th annual meeting of the International Communication Association, San Juan, Puerto Rico, USA.
- Schneier, B. (2012). *Liars and outliers: Enabling the trust that society needs to thrive*. John Wiley & Sons.
- Schwartz, B. (2004). *The paradox of choice: Why more is less*. New York: Ecco.
- Sheehan, K. B., & Hoy, M. G. (2000). Dimensions of privacy concern among online consumers. *Journal of public policy & marketing*, 19(1), 62-73.
- Striphas, T. (2015). Algorithmic culture. *European Journal of Cultural Studies*. Vol. 18(4-5) 395–412.
- Stroud, N. (2011). *Niche news: The politics of news choice*. New York: Oxford University Press.
- Tsesis, A. (2014). The Right to Erasure: Privacy, Data Brokers, and the Indefinite Retention of Data. *Wake Forest L. Rev.*, 49, 433.
- Turow, J. (2011). *The Daily You: How the new advertising industry is defining your identity and your worth*. New Haven, Connecticut: Yale University Press.

- van Cuilenburg, J. (2007). Media diversity, competition and concentration: Concepts and Theories. In Els De Bens (Ed.), *Media between Culture and Commerce*, Vol. 4, 25-54.
- Vissers, S., Hooghe, M., Stolle, D., & Mahéo, V. A. (2011). "The Impact of Mobilization Media on Off-Line and Online Participation: Are Mobilization Effects Medium-Specific?" *Social Science Computer Review*, 30(2), pp. 152-169.
- Weisberg, J., Te'eni, D., & Arman, L. (2011). Past purchase and intention to purchase in e-commerce: The mediation of social presence and trust. *Internet Research*, 21(1), 82-96.
- White, D. M. (1950). The gatekeeper: A case study in the selection of news. *Journalism quarterly*, 27(4), 383-390.
- Wu, J. H., & Wang, S. C. (2005). What drives mobile commerce?: An empirical evaluation of the revised technology acceptance model. *Information & management*, 42(5), 719-729.
- Yan, D. & Chen, L. (2015). An Empirical Study of User Decision Making Behavior in E-Commerce. In F. Nah and C. Tan (Eds.), *HCI in Business, Lecture Notes in Computer Science*, Vol. 9191 (pp. 414-426). Switzerland: Springer International Publishing.
- Yoon, S. J. (2002). The Antecedents and Consequences of Trust in Online-Purchase Decisions. *Journal of Interactive Marketing*, 16 (2), 47-63.
- Zhang, W., Yuan, S., & Wang, J. (2014). Optimal real-time bidding for display advertising. In *Proceedings of the 20th ACM International Conference on Knowledge Discovery and Data Mining (SIGKDD)*, 1077-1086.