

## **COOL OR CREEPY: CONSUMER COMFORT LEVEL WITH SENTIMENT ANALYTICS**

*Matthew North, Utah Valley University, mnorth@uvu.edu*

### **ABSTRACT**

*The ability to use consumer sentiment analytics to identify, categorize and personalize interactions with customers has never been greater. Hardware, software and algorithms now work more seamlessly and at greater speeds to inform businesses of who their customers are and what is important to them. With the power of consumer sentiment analytics comes the promise of more targeted (even personalized) marketing, rapid customer service, more timely order fulfillment, etc. However, danger arises when such data gathering and analytics programs take place in the shadows, outside the awareness of the consumer, and when the results of analytics activities cross the line from being cool to creepy. This paper examines the responses of 56 adult consumers from across the United States in the context of their awareness of data gathering and analysis, their comfort level with these practices, and their perceptions of the results as beneficial and worth the risk (cool), or detrimental and not worth it (creepy). Our results show that the consumers in our study prefer not to have data about them gathered and mined, but if data are gathered, the practice should be disclosed and constrained only to the organization gathering the data.*

**Keywords:** Analytics, Data Mining, Consumer Sentiment, IT Ethics

### **INTRODUCTION**

At IACIS 2014, we presented on our work-in-progress, developing a model and engine for extracting consumer sentiment from large, unstructured data sets. As that project has progressed over the course of the past year, several questions have arisen on topics ranging from algorithm choice to ethics. This paper presents research into one of these questions, which could effectively be summarized as: “Our sentiment analysis engine is really powerful and it’s generating some interesting results, but is it cool, or creepy?” We endeavored to answer this question for ourselves in this present work.

Last year, Peter Leonard asked a similar question more generally while writing in the journal *International Data Privacy Law*. In his paper, he asks if we can “...facilitate socially beneficial uses and applications of Big Data while precluding ‘Big Brother’ or unduly ‘spooky’, ‘creepy’, or otherwise socially or culturally unacceptable Big Data practices?” [13]. Ultimately, these questions are a matter of ethics: Simply because we can design and build a system to extract, interpret and categorize peoples’ feelings, should we? Do customers want us to?

In order to answer our question, we designed a ten-question survey and invited 170 adults from across the United States to respond. Using both quantitative Likert-scale and open-ended qualitative questions, we prompted respondents to share their thoughts about businesses that collect and mine data about them. More specifically, we asked respondents whether they know when companies are gathering data about them, what their comfort level is when companies do gather data, and whether they have ever had experiences where consumer sentiment analytics has led to a creepy feeling or reaction. In this paper we present our findings from that survey, along with general observations and recommendations for researchers using Big Data to examine consumer sentiment.

### **REVIEW OF LITERATURE**

Respect for consumers’ feelings is not a new phenomenon in the age of Big Data and analytics [15, 26], however the ability to amass and rapidly analyze consumer sentiment data sets has never been more extensive and continues to grow [21]. More than a decade and a half ago, researchers were already beginning to recognize and examine the Internet’s potential for gathering immense amounts of data regarding peoples’ feelings, reactions and behavioral patterns [15]. This of course presented exciting opportunities for marketers, product developers and similar professionals, however these opportunities did not, and do not, come without ethical concerns. Dennis Pitta, et al.

had written about the dilemma of highly customized, even personalized data-driven marketing as early as 1998, and in 2003 his team published the following: “The practice holds the promise of very satisfied customers... One of the building blocks of mass customization is knowledge of the customer and his/her complex set of preferences. There is a significant obstacle to gaining this information, namely the growing trend toward consumer privacy” [22]. It is interesting that Pitta and his colleagues characterized consumers’ desire for privacy as an ‘obstacle’, a characterization that the literature shows has evolved significantly in the intervening years.

### **Build Relationships First, Then Gather Data**

Shortly after Pitta’s description of privacy as an obstacle, Chellappa and Sin (2005) suggested that “investments in online personalization may be severely undermined if consumers do not use these services due to privacy concerns” [4]. In their findings they claim “it is of critical importance that vendors understand and evaluate the different values consumers may place in enjoying various types of personalization”, and propose that vendors hoping to successfully use sentiment analysis to inform highly personalized interactions begin by using trust building exercises with their customers. Follow-up research in the years immediately following the Pitta and Chellappa studies has demonstrated an ongoing need for balance between behavior tracking and respect for privacy.

Cooper characterizes this balance, in the context of analytics, as strategic vs. tactical [5]. Many organizations successfully avoid reaching the level of ‘creepy’ with their customers by not over-personalizing. While it may be tempting to ‘micro-focus’, boiling data down to the individual level may be a little too uncomfortable for many consumers, and Cooper suggests that business that do this many be missing tremendous opportunities at a more strategic level that would be better informed through data aggregation and examination of larger categorical trends. By overemphasizing the power of analytics to market to ‘the one’, organizations miss the opportunity to meld databases with market research and competitor intelligence.

Other research seems to disagree with Cooper’s findings and recommendations however. Writing contemporary to Cooper, Hauser stated: “As the company grows, the need for an intimate analytical database becomes even more imperative because it becomes harder to maintain the one to one personal relationship with a larger and ever growing number of clients. Today, many companies are finding that a loss of business is directly related to their lack of intimacy or emotional connection with the customer” [11]. Contrary to Cooper’s ‘big picture analytics’ recommendation, Hauser advocates specificity, exactness, intimacy. For some, the very *use* of the word ‘intimacy’ will cross the line into the realm of creepy. Perhaps both of these are extreme positions.

Evans [8] proposes balance such as that discussed previously in our examination of Chellappa’s work. Evans sees personalization as pragmatic, but potentially creepy if marketers cross customers’ comfort barriers in the effort to create ‘relationships’. As with the word ‘intimacy’, ‘relationships’ (according to Evans) may be just a little too close for some consumers’ comfort. We now live in a data-driven world, but it is not the aggregation of that data that gives consumers cause for discomfort. Rather, it is the way data are used that causes customers to shiver at times. Evans astutely observes that “In an era of self service and disintermediation, marketers are increasingly turning to customer data instead of the tacit understanding of customers they used to possess when interaction was on a more personal basis”; and subsequently he suggests “If marketing reverted to earlier paradigms rather than promising cynical consumers a ‘relationship’, perhaps all parties might actually start to gain more trust in each other. Personal rather than relational interaction would indeed be more pragmatic” [8].

### **Trust is Key**

There is that word again, ‘trust’. Chellappa [4] recommended that highly personalized marketing campaigns are apt to be more successful if founded upon a relationship of trust between company and customers. Perhaps then, it is not the level of personalization that comes across as creepy when using data to discern customer sentiment, but rather, the degree of trust—a lack of which may result in resistance or withdrawal by consumers. Individuals are more likely to accept personalized contact from their vendors if they feel that they are participants in the process. Desouza, et al [6] recommend a customer-centric analytical approach. Acknowledging that technology has radically changed the business-to-consumer interaction, they urge organizations to adopt a view of their customers as empowered partners, as opposed to simply view them as sales targets. Ferguson would agree: “In the consumer world, data tends to flow one way, from the consumer to the database. [Businesses] can also learn a lot about

their...customers by reversing the data stream. [...] The data flow back to the business owner becomes an additional source of understanding” [9]. This concept is strongly reflected in the data we collected for this paper, as we will discuss in our Methodology and Results sections later.

It is apparent from this research that customers are more likely to see our efforts at understanding and interpreting their feeling through data analytics as cool rather than creepy if processes are open and transparent. High-tech personalization cannot be comfortably done in the shadows. In the most current literature, a term for such behind-the-scenes data gathering and mining has emerged: ‘dataveillance’ [7, 28]. In a world now colored by the actions of such figures as Edward Snowden and Julian Assange, this term carries certain connotations for even more passive individuals. There is a secrecy to the term; one that will not contribute the aforementioned trust that is vital to successfully establishing data-driven personalization in customer care. Rubenstein cautions against analytics-in-the-dark tactics: “...regulators should encourage businesses to adopt new business models premised on consumer empowerment...” [23]; a sentiment also reflected by Birnhack, Toch and Hadar: “...privacy design can leave enough leeway for end users to choose how they wish to manage their privacy...” [2]. The ostensible message here is that systems designers can, and should, create mechanisms whereby the consumer is also a participant, and active player in the rules that govern data that are gathered about themselves. Quality personalization and personal privacy should not, and does not have to be an ‘either-or’ proposition, though some organizations often see it as such [14].

### **Customer Data Analytics in the Light**

As organizations succeed in creating this partnership with their customers by conducting their analytical activities in the light, trust will increase and higher levels of personalization will come to be seen as cool rather than creepy. Aguirre, et al. have published recent research demonstrating exactly this point [1]. In their findings, the team reports: “When firms engage in overt information collection, participants exhibit greater click-through intentions in response to more personalized advertisements, in contrast with their reactions when firms collect information covertly. This effect reflects the feelings of vulnerability that consumers experience when firms undertake covert information collection strategies.” Their data demonstrate, in a tangible way, the importance of enrolling customers as partners in the analytical experience, a goal that can only be accomplished in an environment of participation and transparency.

When such an environment does exist, strong positive outcomes have been observed [3, 10]. The Gutiérrez study [10] makes clear that the Internet age has conditioned more people to view themselves, even if only subconsciously, as active participants who desire trust relationships, rather than as passive consumers who are fed messages they are expected to automatically accept. Bragge, et al. support this premise as well: “Although the technological advances allow fancy enhancements in personalization, it is imperative that the context-specific customer attitudes toward online personalization are taken into account by businesses” [3]. Lindstrom advances this same agenda—that consumers should become increasingly active participants as business/data subjects: “While one may take companies to task for their devious use of psychology and technology to ensnare customers...it is ultimately consumers who are responsible for their own behaviors” [16].

### **Building Trust in Today’s Environment**

With the need for trust relationships now established, we must address the question of how this is accomplished. There is much research dedicated to the concept of Customer Relationship Management (CRM), and more emerging research focusing specifically on the use of social media in customer relationship building. We will focus on a sample of this research here. Williams offers a foundation for relationship building with customers, which includes the following attributes: satisfaction, alignment, compatibility, protection against opportunism, communication, empathy, trust and commitment, long-term orientation, effective data management, reciprocity, control, cooperation, productive conflict resolution, adaptability, and personal responsibility and empowerment [29].

When taken together, such relationship building will appear daunting, but this foundations is both necessary and nurturing if we want our analytical approaches to personalization to succeed. Miles suggests using a measured and incremental approach to building upon Williams’ proposed foundation—a proverbial ‘how to you eat an elephant’ recommendation [17]. Min, et al. [18], Yee Liau and Pei Tan [30] and Sebastian [25] each propose that companies focus in their area of specific product or service offering. In other words, do not try to build an overarching or all-encompassing relationship with your customer, but rather, focus on creating that connection in the area where you

and they interact. For example, it is not necessary to spend time and resources establishing a relationship of fiduciary trust if your interactions do not include a monetary component. The focus should be on meeting the consumer where the business's offerings and the customer's needs intersect. This is much simpler, more focused, and more likely to result in a relationship of trust.

As more and more people turn to participation in social media as an integral part of their daily lives, so too have corporations sought out uses for these technologies to connect with their customers. In the mid-1990s, we observed a marked shift in marketing and promotion where more and more impressions included a web site address. In the new millennium, this phenomenon has evolved to include (and even emphasize) a Twitter handle, a YouTube channel, an Instagram account, a Facebook page, etc. [24]. Social media has assumed an entrenched place in modern life, communication, and interaction, and as such will continue to play a stronger and more central role in the relationships organizations have between themselves and their customers. Nadeem asserted in 2012 that: "Social Media (SM) has become an undeniable influence on customer care [...and] engaging customers through social communities is having far-reaching effects on customer loyalty" [19]. Writing again on the topic this year (2015), Nadeem followed up with research indicating that companies, at present, are not succeeding very well at integrating their social media usage with their overall strategies, negatively impacting the trust and relationship they need to establish with consumers [20]. Social media tools are not only a means of connecting and communicating with consumers on a personal level, they are also inherently powerful data gathering mechanisms. As has already been established however, using them as a data source must be done openly, and as part of an integrated strategy, or they risk becoming a detriment to trust relationships, rather than an asset.

### **Know Thyself, Know Thy Audience**

The research cited thus far suggests that consumers do want and appreciate personalized interaction, but only within their comfort level—a comfort level that is largely established upon a relationship of trust. Such relationships are unlikely to be successfully created purely through the use of automation [27]. Even when analytics can point us in the direction of attributes in a data set that indicate what consumers value, it is the analyst, the human component, that must determine meaning and appropriate application. Understanding customers' values, motivations and expectations can be informed through data analytics, but if we apply what we learn in ways that feeling invasive to those peoples' lives, we risk undermining ourselves and our efforts to serve those customers. Jiao, Zhou, et al. wrote specifically on this matter in 2007 [12], and again in 2013 with updated observations and recommendations [31]: "Mass personalization aspires to assist customers in making better informed decisions, and to the largest extent, to anticipate customer satisfaction and adapt to customer delight." To accomplish this, the research agrees with the premise we have sought to establish in this literature review: Consumer sentiment analytics is perceived as cool, rather than creepy, when it is conducted in open view of the consumer, even enlisting him/her as an active participant, with an end to better addressing his/her needs.

Based upon this literature, we felt it was extremely important to include a component in our own research agenda that will inform our usage of data and consumer sentiment analysis. To return to our previously stated research question, we want to be sure that we can answer: "Simply because we can design and build a system to extract, interpret and categorize peoples' feelings, should we? Do customers want us to?"

### **RESEARCH METHODOLOGY**

With our research question stated and reiterated above, contextualized in the research and literature we have cited and discussed, we propose the following hypotheses for this paper:

H<sub>1</sub>: Consumers claim to be aware of when organizations gather data about them.

Sub-H<sub>1</sub>: There is no difference in consumer awareness on H<sub>1</sub> between online and bricks-and-mortar establishments.

H<sub>2</sub>: Consumers are comfortable with organizations gathering data about them.

Sub-H<sub>2</sub>: There is no difference in consumer comfort on H<sub>2</sub> between online and bricks-and-mortar establishments.

H<sub>3</sub>: Consumers are comfortable with interactions that result from organizations using data about them.

To address these hypotheses, we developed a 10 question survey and extended 170 invitations to complete it to adults across the United States. The survey was administered online, and results were gathered and analyzed electronically. We intentionally kept the survey brief in an attempt to maximize participation, and we did not offer any incentive or enticement for survey completion in order to avoid unduly influencing or skewing the results.

Respondents provided an indication of their age by selecting from one of five ranges: 20 years or less; 21-30, 31-40, 41-50, or 50+. They also provided their sex as Female, Male or Other, with the ability to specify the meaning of Other. We then specifically asked survey participants, using a five point Likert scale, how often they feel they are aware that a company is gathering data about them, both in traditional ‘bricks-and-mortar’ settings, and online. We also asked how comfortable respondents were with such data gathering activities, both in person and online. Further, we asked if participants had ever received marketing or other personalized contact from a company that was clearly the result of data analytics, and was so personalized that the recipient considered the experience to be ‘creepy’, and if so how often. Finally, we presented the participants with an opportunity to provide free-form, qualitative comments about their beliefs, observations or experiences with consumer data gathering and analysis. A combination of descriptive statistics with visualizations, *t*-tests, Pearson correlation, and linear regression were used to analyze the data.

## RESULTS

Of the 170 invitations to participate in the survey, 56 responses were received, yielding a 33% response rate. While we had hoped for more, we felt strongly that not influencing or coercing participation was important to study integrity. We recognize a small sample size as a limitation of this study, while asserting that our data represent real consumer opinions and are therefore still valid, if only on a limited scale. Our data distribution was extremely normal, with standard deviations for all Likert-based questions very close to one. One response was incomplete, having one question left blank. Because of our normal distribution, we felt comfortable using linear regression to interpolate the missing value. Our regressions model was highly significant ( $F: .6^{e-7}$ ) and highly predictive (Multiple R: 80%,  $\alpha=.05$ ), yielding a 3 (neutral) on the question about comfort level with bricks-and-mortar organizations gathering consumer data.

### Demographics

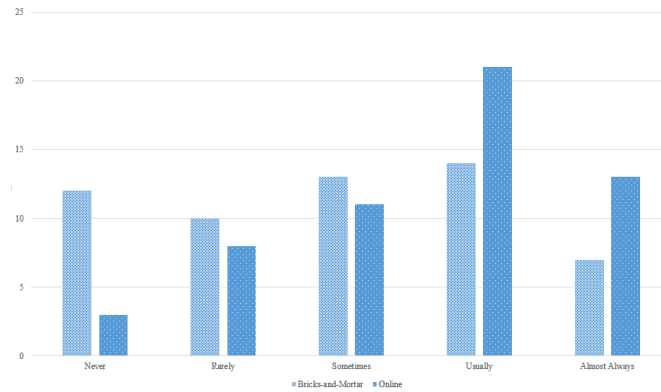
Twenty-three respondents were women, 32 were men, and one self-identified as transgender. Age was distributed across all five age categories, with slightly more than half (64%) under the age of 31. Upon conducting a Pearson correlation, neither Age nor Sex correlated significantly with any other variable in our data set. Thus, we will not break down any further analysis in this paper along these demographic variables.

**Table 1.** Demographics

Sex	<21	21-30	31-40	41-50	50+	Total
Female	6	9	5	1	2	23
Male	10	10	5	6	1	32
Transgender	0	1	0	0	0	1
<b>Total</b>	<b>16</b>	<b>20</b>	<b>10</b>	<b>7</b>	<b>3</b>	<b>56</b>

### Awareness

In our literature review, we extensively discussed research which underscores the importance of conducting customer data analytics in the light of day. Our hypothesis H<sub>1</sub> supposed that consumers are generally aware of when companies are gathering data about them (or believe they are aware), and further supposed (Sub-H<sub>1</sub>) that there was no difference in claimed awareness between online and bricks-and-mortar organizations.

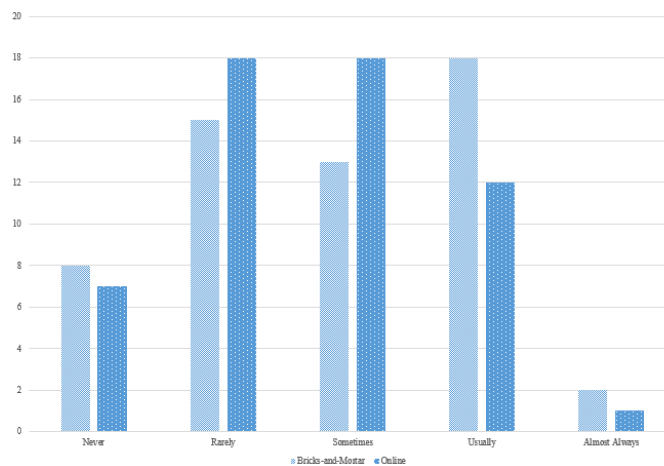


**Figure 1.** Claimed Consumer Awareness of Data Gathering and Analytics

Our respondents indicated that 61% of the time in a bricks-and-mortar setting, and 80% of the time when online, they are aware of when data are being gathered about them. Taken together, 71% of our participants were either ‘Sometimes’, ‘Usually’, or ‘Almost Always’ aware of companies’ data gathering activities. Thus we fail to reject  $H_1$ , and accept that people are generally aware of when consumer data about them is being collected and used. Sub- $H_1$  further theorized that there would be no significant difference between in-person and online awareness. A two-tailed  $t$ -test on these two questions yielded a  $p$ -value of .002 ( $\alpha=.05$ ,  $t$ -Stat: -2.939), therefore we reject the null hypothesis and accept the alternate: there is a statistically significant difference between respondents’ awareness of data gathering online versus in bricks-and-mortar environments. We find this curious: one might assume that a consumer would be more likely to know if data gathering and analysis mechanisms were in place when standing in a store, for example, than when browsing on a website. One possible explanation is that respondents simply assume that when interacting with a company online, data are being gathered, and so they responded on the survey indicated higher degrees of awareness. This is a question worthy of further investigation. In addition to the statistically significant difference between online and in-person data gathering/analytics awareness, there was also a statistically significant Pearson correlation between these two as well. As awareness in bricks-and-mortar situations rises or falls, so too does awareness in the online setting ( $p=.59$ ). This is perhaps not as surprising as the gap between the two, since those who are more aware in general, are likely to be more aware in one environment or the other.

### Comfort

Our next hypothesis theorized that customers are comfortable when data about them is gathered and analyzed. We also further hypothesized that there would be no difference between online and in-person settings related to levels of comfort.



**Figure 2.** Consumer Comfort with Data Gathering and Analytics

Figure 2 illustrates that respondents' comfort level with data collection and analysis is generally lower than is their awareness. Only two responses indicated they are 'Almost Always' comfortable with bricks-and-mortar establishments gathering and analyzing data about them, and only one felt that way when online. By contrast, participants indicated that 64% of the time for in-person and 77% of the time for online, they are either 'Never', 'Rarely', or 'Sometimes' comfortable with data gathering by the companies they patronize. Even when removing the 'Sometimes' values, discomfort percentages remain above 40% (41% and 45% respectively). Based on these indicators, we reject  $H_2$ , accepting that customers are generally *not* comfortable with businesses gathering and analyzing data relating to their consumer behavior. Similar to Sub- $H_1$ , Sub- $H_2$  supposed that there would be no significant difference between in-person and online comfort. A two-tailed  $t$ -test on these two questions yielded a  $p$ -value of .431 ( $\alpha=.05$ ,  $t$ -Stat: .789), causing us to fail to reject the null hypothesis: when it comes to data gathering, there is no statistically significant difference between respondents' comfort online versus in bricks-and-mortar. Here, our Pearson correlation strongly supports our findings: as comfort in bricks-and-mortar situations rises or falls, so too does awareness in the online setting ( $p=.76$ ). This appears to support the overall observation that in our sample of survey participants, people simply aren't very comfortable with the way businesses collect and analyze data about them.

### What's the Verdict: Cool or Creepy?

Our data establish that people generally feel aware that companies are gathering data for consumer analytics, and they (the consumers) are usually not very comfortable with it. After asking the questions that helps us to arrive at this conclusion, we asked respondents four additional questions to potentially inform why they feel the way they do. These were intended to help us answer the 'Cool or Creepy' question more directly. Our hypothesis  $H_3$  suggests that people feel comfortable with *how* organizations use data gathered about consumers.

Here we acknowledge that consumers do not know all of the ways companies use data that are gathered. Despite this limitation, we did ask some questions about general user experience that can inform customer understanding of how data are used once collected. The first of our questions regarding how data are used asked how often respondents had received or experienced a tailored interaction with a company that was so personalized it came across as 'creepy' [13]. Overwhelmingly, the answer is 'not very much', with 75% responding 'Never' or 'Rarely' (11% responded 'Frequently' or 'Very Often'). The next of these questions asked if respondents would be more comfortable with data gathering and consumer analytics if the activity resulted in them receiving a better deal or a higher level of service. Perhaps not surprisingly, 32% said they would (answering 'Agree' or 'Strongly Agree' to the question). Regarding sharing of gathered data with third-parties or other organizations, respondents overwhelmingly either 'Strongly Disagreed' or 'Disagreed' with the practice (73%), underscoring the message that most do not want data about their consumer behavior collected and passed around like a commodity. Figure 3 summarizes the responses to these three questions about how consumer data is used. The same Likert scale was used for all three, although on the question about creepy experiences, the labels ranged from Never (1) to Very Often (5); while on the questions about getting a better deal and businesses sharing data with other parties, the labels ranged from Strongly Disagree (1) to Strongly Agree (5).

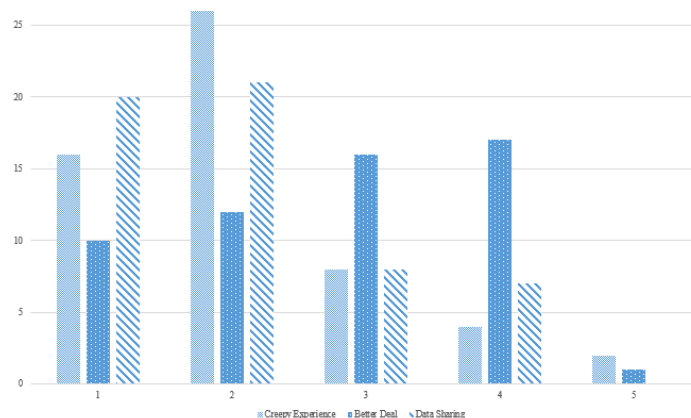


Figure 3. Consumer Opinions on How Data Get Used in Analytics

As evidenced in this chart, most of our respondents have not had many experiences with personalized business contact that they considered to be creepy. If we use the question about analytics leading to better deals or higher quality service to indicate a cool factor, there is some evidence that consumers will accept a level of personal data analytics, however the majority have still indicated that they would prefer not to be data mined, even if it might lead to a better customer experience. Finally, the participants in our study have indicated resoundingly that if businesses are going to gather and analyze data, they ought not to share it with others. In analyzing these data statistically, a significant Pearson correlation coefficient of .66 between the 'better deal' question and the 'third-party sharing' question indicates that the more a respondent is willing to accept consumer analytics to get a better deal, the more they are also willing to let that data be shared. In using linear regression, we find that we can use the 'awareness' and 'comfort' questions as predictors of our respondents' experiences with creepy interactions ( $F=.012$ , Multiple  $R=46\%$ ,  $\alpha=.05$ ) and of their willingness to get a better deal through analytics ( $F=.007$ , Multiple  $R=49\%$ ,  $\alpha=.05$ ). Though these figures are statistically significant, the predictive power in our particular data set is relatively weak. Thus, we can describe our participants' preferences relative to consumer data analytics, however we would also be wise to limit inferences to larger populations based on this study.

In the final question, where we invited respondents to share qualitative observations, we received a number of comments that underscore our quantitative findings. The quotes offered here characterize the main points we observed in the data set, some of which were made multiple times:

- "I'm completely okay with general market and demographic research, however I am not okay with selling data, or overly-personalized marketing emails. I know that my information is being gathered nearly all the time, as I generally assume that if there's information available, it's being gathered."
- "I like the way online does it. They keep records of what I buy without bothering me about buying other stuff I don't necessarily want or need."
- "Retailers shouldn't gather/track data concerning my shopping habits or browsing patterns."
- "I believe they should always notify you and allow you to opt-out of data mining activities."
- "Retailers use data to improve their service to their customers, and I want to be treated the best I can be, so if gathering and mining data about me improves how I'm treated (timeliness, accuracy, better pricing), then I am in favor of it. That being said, I want retailers to tell me that they're gathering data, what data they're gathering, and how they're protecting my privacy. I also expect them not to share it with other companies, even though I know many of them do."
- "I don't have problem about retailers gathering data that pertain to my transactions. I just don't want them to use my data to create new shopping habits that I didn't have before."
- "I understand that the data mining happens and is a part of doing business, but I don't like the idea of them selling that information to another company especially if I don't like doing business with the other company. If they want to use my information to improve their customer service, and product availability that is fine, as long as they are using it and not someone else."
- "Attempts should be made (by humans, not machines) to determine when information is so sensitive that it should not be used for data mining purposes or to send further information to me based on what they have gathered."
- "I really don't have a problem with retailers gathering information as long as they keep that information to themselves. My discomfort increases when they sell that information. I also don't like getting international calls with questions about me."
- "What creeps me most is after I've done eye shopping on a specific website, the same item I clicked keeps following me. If they want to be smart, they shouldn't do that."
- "Being contacted as a result of data mining activity can motivate the same feelings in individuals that are experienced during telemarketing phone calls. Perhaps consumers can be given a choice by establishments to say yes or no to having their data collecting for marketing purposes. I feel that I would not mind the establishment gathering my data as long as it was not sold to any other organization without my knowledge. In fact, establishments should offer this choice to consumers as well."

The general theme that can be taken away from these comments ties back into our  $H_3$  research question: "Consumers are comfortable with interactions that result from organizations using data about them." The correlational and descriptive statistics offered in this section, along with these qualitative comments, lead us to reject



H<sub>3</sub>. Gratefully, few of our participants reports experiencing a creepy interaction as a result of consumer sentiment analysis, however the overwhelming message sent through responses to the ‘better deal’, ‘third-party sharing’ and open-ended question is that consumers, at least in this study, are *not* comfortable with the way data and analytics are being used.

### CONCLUSIONS

Our research in this paper has shown that consumers generally feel aware of data gathering activities (or at least that they claim to be aware); that they prefer that organizations not gather data about their buying behaviors; and that if they do, those organizations both disclose the practice and keep collected information private. While not rising to the level of being creeped out by businesses conducting consumer sentiment analytics, the respondents indicate a desire to be both informed and protected. Our literature review clearly indicated that our findings are consistent with existing and current research in the field—people are more apt to accept being data mined if they trust the organization doing the mining and feel like a partner, rather than like a subject, in the process. One question we might have asked in our survey but did not would be whether or not respondents generally trust the organizations with which they do business in person and online. A correlation between the ‘awareness’ questions, the ‘comfort’ questions, and this hypothetical ‘trust’ question might be interesting and informative in light of the theme of trust that came out in our literature review. Although our respondents did not strongly indicate that data-driven personalization has been creepy with much regularity, additional inquiry into why consumers might feel creeped out could be an interesting avenue for further study.

Consumer sentiment analytics is not new, but has been radically re-invented in the past decade and a half, thanks to faster, more powerful and more accessible technology. Perhaps the greatest take-away from this paper is for businesses to remember that personalization through technology must connect to a person, not a profile. If the former is accomplished, then the analytical technology is more likely to be perceived as cool and therefore appreciated by the consumer; if the latter is realized, that personalization may just come off as creepy.

### REFERENCES

1. Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness. *Journal of Retailing*, 91(1), 34-49.
2. Birnhack, M., Toch, E., & Hadar, I. (2014). Privacy Mindset, Technological Mindset. *Jurimetrics*, 55(1), 55-114.
3. Bragge, J., Sunnika, A., & Kallio, H. (2012). An Exploratory Study on Customer Responses to Personalized Banner Messages in the Online Banking Context. *Journal of Information Technology Theory and Application*, 13(3), 5-18.
4. 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.
5. Cooper, T. (2006). Enhancing Insight Discovery by Balancing the Focus of Analytics between Strategic and Tactical Levels. *Journal of Database Marketing & Customer Strategy Management*, 13(4), 261-270.
6. Desouza, K. C., Awazu, Y., Jha, S., Dombrowski, C., & Papagari, S. (2008). Customer-driven Innovation. *Research Technology Management*, 51(3), 35-44.
7. Esposti, S. D. (2014). When Big Data Meets Dataveillance: The hidden side of analytics. *Surveillance & Society*, 12(2), 209-225.
8. Evans, M. (2003). The Relational Oxymoron and Personalisation Pragmatism. *The Journal of Consumer Marketing*, 20(7), 665-685.
9. Ferguson, R. (2009). The Consumer Inside: At its heart, all marketing speaks to human beings. *The Journal of Consumer Marketing*, 26(3), 214-218.
10. Gutiérrez, S. S., & Izquierdo, C. C. (2010). Los Determinantes de la Confianza del Comprador Online. Comparación con el caso de subasta. *Cuadernos de Gestión*, 10, 43-61.
11. Hauser, W. J. (2007). Marketing Analytics: The evolution of marketing research in the twenty-first century. *Direct Marketing*, 1(1), 38-54.

12. Jiao, R. J., Xu, Q., Du, J., Zhang, Y., & Helander, M. (2007). Analytical Affective Design with Ambient Intelligence for Mass Customization and Personalization. *International Journal of Flexible Manufacturing Systems*, 19(4), 570-595.
13. Leonard, P. (2014). Customer Data Analytics: Privacy settings for 'Big Data' business. *International Data Privacy Law*, 4(1), 53-68.
14. Li, T., & Unger, T. (2012). Willing to Pay for Quality Personalization? Trade-off between quality and privacy. *European Journal of Information Systems*, 21(6), 621-642.
15. Limehouse, D. (1999). Know your Customer. *Work Study*, 48(3), 100-102.
16. Lindstrom, M. (2012). Brandwashed: Tricks Companies Use to Manipulate Our Minds and Persuade Us to Buy. *Mankind Quarterly*, 52(3/4), 428-438.
17. Miles, D. A. (2014). Measuring Customer Behavior and Profitability: Using Marketing Analytics to Examine Customer and Marketing Behavioral Patterns in Business Ventures. *Academy of Marketing Studies Journal*, 18(1), 141-165.
18. Min, H., Hyesung, M., & Emam, A. (2002). A Data Mining Approach to Developing the Profiles of Hotel Customers. *International Journal of Contemporary Hospitality Management*, 14(6), 274-285.
19. Nadeem, M. (2012). Social Customer Relationship Management (SCRM): How Connecting Social Analytics to Business Analytics Enhances Customer Care and Loyalty? *International Journal of Business and Social Science*, 3(21), na.
20. Nadeem, M. (2015). Customer Loyalty at Any Cost: Why Is Social Media Poorly Integrated With Marketing Strategy? *International Journal of Academic Research in Business and Social Sciences*, 5(2), 78-96.
21. Phillips-Wren, G., & Hoskisson, A. (2015). An Analytical Journey Towards Big Data. *Journal of Decision Systems*, 24(1), 87-102.
22. Pitta, D. A., Franzak, F., & Laric, M. (2003). Privacy and One-to-one Marketing: Resolving the conflict. *The Journal of Consumer Marketing*, 20(7), 616-628.
23. Rubenstein, I. S. (2013). Big Data: The End of Privacy or a New Beginning? *International Data Privacy Law*, 3(2), 74-87.
24. Rust, R. T., & Huang, M.-h. (2014). The Service Revolution and the Transformation of Marketing Science. *Marketing Science*, 33(2), 206-221.
25. Sebastian, T. (2010). Deriving Customer Loyalty and Its Determinants from Online Reviews using Support Vector Machine. *Journal of Internet Business*, 8, 61-80.
26. Tavani, H. T. (1999). Informational Privacy, Data Mining, and the Internet. *Ethics and Information Technology*, 1(2), 137-145.
27. Tseng, T.-L., Huang, C.-c., Ho, J. C., Pannathur, A., & Zhu, Y. (2008). Autonomous Decision Making in Customer Relationship Management: A Data Mining Approach. *IIE Annual Conference. Proceedings*, 721-726.
28. van Dijck, J. (2014). Datafication, Dataism and Dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, 12(2), 197-208.
29. Williams, K. C. (2012). Core Qualities of Successful Marketing Relationships. *Journal of Management and Marketing Research*, 10, 1-29.
30. Yee Liau, B., & Pei Tan, P. (2014). Gaining Customer Knowledge in Low Cost Airlines through Text Mining. *Industrial Management + Data Systems*, 114(9), 1344-1359.
31. Zhou, F., Ji, Y., & Jiao, R. J. (2013). Affective and Cognitive Design for Mass Personalization: Status and prospect. *Journal of Intelligent Manufacturing*, 24(5), 1047-1069.