Media and Communication
Open Access Journal | ISSN: 2183-2439
Volume 6, Issue 3 (2018)

Multidisciplinary Studies in Media and Communication

Editors

Epp Lauk and Raul Reis

COGITATIO

Media and Communication, 2018, Volume 6, Issue 3
Multidisciplinary Studies in Media and Communication

Published by Cogitatio Press
Rua Fialho de Almeida 14, 2º Esq.,
1070-129 Lisbon
Portugal

Available online at: www.cogitatiopress.com/mediaandcommunication

Table of Contents

COGITATIO

Media and Communication (ISSN: 2183–2439)
2018, Volume 6, Issue 3, Pages 60–68
DOI: 10.17645/mac.v6i3.1205

Article

Personal Power and Agency When Dealing with Interactive Voice Response Systems and Alternative Modalities

Jill Walsh, Brittany Leigh Andersen *, James E. Katz and Jacob Groshek

Division of Emerging Media Studies, Boston University, Boston, MA 02215 USA; E-Mails: jillw@bu.edu (J.W.), bla@bu.edu (B.A.), katz2020@bu.edu (J.K.), jgroshek@bu.edu (J.G.)

* Corresponding author

Submitted: 1 October 2017 | Accepted: 16 March 2018 | Published: 25 September 2018

Abstract
In summer 2015, we conducted an exploratory study of how people in the U.S. use and respond to robot-like systems in order to achieve their needs through mediated customer service interfaces. To understand this process, we carried out three focus groups sessions along with 50 in-depth interviews. Strikingly we found that people perceive (correctly or not) that interactive voice response customer service technology is set up to deter them from pursuing further contact. And yet, for the most part, people were unwilling to simply give up on the goals that motivated their initial contact. Consequently, they had to innovate ways to communicate with the automated systems that essentially serve as gatekeepers to their desired ends. These results have implications for communication theory and system design, especially since these systems will be increasingly presented to consumers as social media affordances evolve.

Keywords
computer mediated communication (CMC); interactive voice response systems (IVRs); media equation theory; power in communication; social robots; theory of mind

Issue
This article is part of a multidisciplinary issue of *Media and Communication*, edited by Epp Lauk (University of Jyväskylä, Finland) and Raul Reis (Emerson College, USA).

1. Introduction

Interactive voice response systems (IVRs) and other robot-like technology are used widely by corporations and other institutions as a means to address their clients' service and information needs. IVRs are automated systems that allow a user to interact with a robot remotely, via the phone or computer. IVR robots are often used in customer service to moderate standardized and repetitive inquiries, such as with appointment scheduling, refilling prescriptions, insurance questions, or collecting bank account information. IVR systems prompt users with specific questions and direct them to different menus options based on the voice or keypad input. As such, IVRs are set up as gatekeepers to information; users must answer questions and/or provide information in order to get to the help they seek. In an effort to further understand user interactions with IVR technology, we conducted an exploratory study of U.S. uses of IVR and related systems in the fall of 2015 to get deeper insight into how people communicate with power in this institutional context. Though not often a subject of academic study, this process is one that affects millions of people daily in the U.S., and as modern communication technologies spread, it will increasingly do so around the world.

Providing clients with such inbound communication opportunities, and dealing with the demands that often follow in the wake of them, are resource consumptive for organizations, and thus they have an incentive to automate and cheaply channel such services. Specifically, from an institutional perspective, using IVRs as gatekeepers makes financial sense as it reduces the need for paid employees to manage phone lines, and in theory, the wait-time for the client. Institutions can also argue that

IVRs deliver in many cases superior customer service, leading to more satisfied clients. As a result, IVRs are almost a universal component of the client service experience and can be found in a gamut of industries, organizations, and institutions. Yet unsurprisingly, only 3% of respondents in our 2015 study reported that they liked using IVRs as part of their client service experience (Center for Research on the Information Society, 2015). Clients are thus faced with a dilemma; they need help from the institutions but dislike IVR interactions, yet despite their dislike, they are routinely forced to engage with the system to achieve their desired ends.

This conundrum in essence allows us to confront a major theme of scholarship concerning politics and communication technology, namely communicating with power. When clients contact agencies with a problem, either to request information or to address a need, they are cast in the position of a supplicant as the institution has the power to take action or refrain from doing so. Knowledge is also unequal as the institution has thorough information about what can be done to help the client and in many cases a trove of personal information about the client. Though the communicational power asymmetries are clear under the circumstances, they need not be exaggerated either. For it is also the case, based on research presented here and that of others, that people are resourceful in negotiating a set of institutional hurdles to pursue their own interests. Specifically, to manage dealing with restricted access to institutional resources that is forced upon them through automatic systems, people invoke strategic decision-making in response to the IVR gatekeeping regimes they encounter. The extant research literature covers some structured lab experiments (e.g., Groom et al., 2011) and older public opinion data (Katz, Aspden, & Reich, 1997) about how people manage client service. Our search of the literature reveals no instances of a nuanced interpretation of these processes despite the fact that millions of people engage in them every day.

Our purpose therefore is to complement the research literature by exploring the innovative ways people use to manage their relatively and sometimes nearly absolutely powerless position in communicating via IVR client service interfaces. This work is important because in addition to explaining the client service experience it also serves as an interesting example of how people approach technological structures of power and express their human agency when dealing with automatic entities that can often render them essentially powerless. Thus it connects with another set of literature, in this case massive, concerning individual lives in relationship to institutional power structures (e.g., Castells, 2009; Foucault, 1984) but due to space limitations can only be touched upon.

2. Powerful Communications in Review

Ultimately, customer service contacts are negotiations between unequal players with asymmetrical knowledge. Foucault (1984) argued that "power is everywhere; not because it embraces everything, but because it comes from everywhere" (p. 93). He goes on to say that power is not "acquired, seized, or shared" but rather something that is "exercised" in unequal relationships (p. 94). Castells (2009) is especially helpful in understanding the ways in which clients negotiate this process. Like Foucault, Castells believes that power is everywhere, but he defines it as the "relational capacity that enables a social actor to influence asymmetrically the decisions of other social actor(s) in ways that favor the empowered actor's will, interest and values" (2009, p. 10). Through these mechanisms, power is embedded in relationships rather than innate characteristics of the actors or groups (van Dijk, 2010).

This is a critical component of the customer service experience as it is the relationship between the IVR and client, in particular the IVRs' location as gatekeeper in this process, that gives it such power. We simply cannot receive a response to our question without the IVR. Castells (2009) uses the term network power to refer to a power relationship in networked communication in which protocols dictate the process. Network power is critical customer service interactions as clients are beholden to the protocols devised by the programmers and are not able to vary their communication style. This is especially important as the protocols are specifically designed to limit client interaction and in essence make the client go away.

2.1. Cooperating with Customer Service Systems

The notion that people respond socially to computers has been well documented in laboratory settings (Halpern & Katz, 2013; Nass & Brave, 2005; Nass, Fogg, & Moon, 1996; Reeves & Nass, 1996) and through survey research (Katz & Halpern, 2013). As the brain cannot make distinctions between speaking with a robot or human, people end up "applying the same rules and shortcuts that they use when interacting with people" (Nass & Brave, 2005, p. 4). Indeed, the social response appears to be automatic in people, triggered simply by a voice speaking to us, even if it is a very synthetic voice (Nass & Brave, 2005). This would suggest that people interacting with IVRs for client service help will be inclined to follow social norms and engage in polite interactions. Of further interest for this work, it appears that when subjects are told they are working with the computer, they begin to "affiliate with the computer as a team" (Nass et al., 1996). One could imagine a situation where client service may feel like teamwork; the client and IVR do have to work together to some extent to ensure there is resolution to the problem. Interestingly, when users are asked about this, they deny these social behaviors (Nass & Moon, 2000).

Power is embedded in the client service experience in an interesting way that may add to the need to engage in impression management work (Goffman, 1959) with the IVRs. According to the survey conducted by the

Center for Research on the Information Society (2015), 90% of respondents want to speak to a live client service agent when they begin their experience. As IVRs are the gatekeepers to the live agent, people need to engage the technology to get to the agent they seek. Work by Fairhurst and Chandler (1989) and Tjosvold (1984) indicate the value of cooperation and smooth social interactions for affecting positive outcomes in negotiations. This data suggest that it may be in the client's best interest to employ impression management work to engage in social norms of politeness with the IVR in order to curry favor as they would with a person. Given that they are in a position of limited power, individuals can use social norms as a means to create power in these interactions. This led to the following research question:

> RQ1: Will users' behavior with IVRs change based on their perceived amount of power?

There are significant communicational interaction ramifications for interacting with IVRs in this manner. First, we would assume that if people unconsciously view their experiences with IVRs as a social interaction, they will then concern themselves with the presentation of self and impression management work that is a critical component of social interactions (Goffman, 1959). When research suggests that people are polite to computers or robots (Reeves & Nass, 1996), it is of course because of social norms about how we treat people as the authors indicate, but it is also due to our concerns about self-presentation. We perform on the public front stage (Goffman, 1959) as we believe is expected of us to ensure that we are regarded positively by others. If the work highlighted above by Nass and his colleagues is true, then the interactions with IVRs, automated systems like robots, should provoke the same internal desires for impression management that we see in these offline interactions highlighted by Goffman. Thus, the following hypothesis is drawn:

> H1: Users interacting with IVRs for client service help will be inclined to follow social norms and engage in polite interactions.

2.2. The Theory of Mind in Computer Interactions

One of the ways that people manage social interactions is to employ theory of mind (Premack & Woodruff, 1978) techniques, which enable us to assign certain qualities and attributes to others. Using theory of mind helps us to feel that we can understand the feelings, knowledge, and perspective of others. We do this intuitively, as we "naturally explain people's behavior on the basis of their minds; their knowledge, their beliefs, and their desires, and we know that when there is a conflict between belief and reality it is the person's belief, not the reality that will determine their behavior" (Frith & Frith, 2005, p. R644). This is helpful as it enables us to understand and predict the other's behavior and manage the uncertainties associated with social interactions.

Although we have the innate ability to engage theory of mind, we are not all universally adept at it (Lyons, Caldwell, & Shultz, 2010). Importantly though, people believe that they are quite good at it and for the most part we enter into our interactions assuming we can understand the other's mental state. Some have argued that the theory of mind ability may give rise to greater levels of manipulation and intentional deception of others (Frith & Frith, 2005); the idea here is that if I have a sense of what you know, I can decide what to share, in good faith or not. However, others have suggested that those with high theory of mind ability are actually less likely to engage in manipulation of others as they tend to score higher on empathy (Lyons et al., 2010).

There are some studies that have considered the ways that people engage theory of mind with computers (Kramar, von der Putten, & Eimler, 2012; Rilling, Sanfey, Aronson, Nystrom, & Cohen, 2004; Thomas, 2001), but it is surprising how little work has been done in this area given people's likelihood to engage with machines socially. The work of Rilling et al. (2004) suggests that interactions with computer partners activate parts of the brain associated with theory of mind, although the effect is not as strong as it is for interactions with human partners. Nonetheless, Functional magnetic resonance imaging (fMRI) data suggests that computer interactions can stimulate some theory of mind behaviors (Rilling et al., 2004). We posit that people use theory of mind as an important tool to negotiate their customer service experiences via IVRs:

> H2: Users will employ theory of mind when engaging with IVR technology.

3. Methods

As this was an exploratory study, we conducted structured interviews and focus groups in order to gain insight into our research question and hypotheses. Our team aimed to see how people managed client service contacts given their limited power to bypass the IVR gatekeepers. Fifty interviews were conducted either in person or phone by the primary investigator over the course of a two-month period in the summer of 2015. The goal of the interviews was to ask about participants' general attitudes towards client service and IVR generally, as well as their specific experiences with certain client service platforms such as speaking with a live agent, using live chat, etc. IVRs, in the form of automated virtual agents and other voice controlled assistants like Siri, were discussed in-depth in order to assess knowledge and comfort with the technology as well as to understand how they use, or strategically opt out of using, the IVR technology. IVRs, robots, and computers were not explained to participants via explicit definitions. Participants did not ask for clarification or definitions of these terms during

the interviews or focus groups, and they often used IVRs, robots, and automated computer systems synonymously when expressing their experiences. They were provided up front with the consistent idea of IVRs: in the realm of customer service, they would have to interact with a predefined or automated system. This was all with the general framing that they are interacting with non-human devices. Interviews ranged from 15 minutes to an hour, with the average about 25 minutes. All interviews were audio recorded.

Three focus groups were conducted in order to test these topics in a group setting. Focus groups were an integral component of this work because they provide insight into "not only what people think but how they think and why" (Kitzinger, 1995, p. 299). The focus group leader did not need to ask many questions to stimulate intense discussion among participants: everyone had memorable client service experiences that they wanted to share. This afforded the opportunity to listen for the moments of consensus in their experiences. Focus groups allow "co-construction of meaning" (Tiggemann, Gardiner, & Slater, 2000, p. 646), and opportunities for participants to question each other (Neuman, 2006), both of which provided deep insight into this topic. For example, in each group there was a significant debate about the experience of talking to automated systems during these calls, and particularly whether the IVR robot can or should express sympathy. The discussions in the groups were far richer than those in the interviews because people could question each other's assumptions and understandings. All groups were audio recorded and ranged between 45–75 minutes in length.

3.1. Demographics and Recruitment

Participants for both the focus groups and the interviews were recruited via snowball sampling and using the researchers' own personal networks. Despite its obvious weakness, this recruitment strategy nonetheless follows early work on the internet that recruited through local personal networks (Kraut et al., 1998). However for our purposes we advertised the study via several researchers' social media platforms. No incentives were offered for participation, although food was provided for focus group participants. Every effort was made to recruit a representative sampling of people based on age, race/ethnicity, and gender. Granovetter's (1973) concept of the strength of weak ties was employed to obtain access to new opinions because weak ties offer entrée to people in different positions in the social network, who will most likely possess different information and ideas.

Focus group recruitment was targeted so that the groups were clustered by age and degrees of self-assessed skill and comfort with technology. The first group (n = 6) comprised young (ages 19–29) heavy tech users and had four women and two men; the second group (n = 7) was older (60s and 70s) and involved low to medium technology users, with four men and three women. The final group (n = 8) included four men and four women in their 30s and 40s and was the only group that included some mix of technology use; there were some professionals who were deeply connected to their technology and early adopters of new features and devices. Others were more ambivalent about technology and/or relied on only a few key features such as text messaging and email.

3.2. Data Analysis

The data collection and analysis was rooted in the grounded theory approach (Glaser & Strauss, 1967). We allowed the theory to emerge from the data and focused on extensive use of note taking and memo writing during the data collection and analysis processes. The primary investigator took field notes during all interviews and focus groups. After each interview or focus group she wrote memos (Covan, 2007; Glaser & Strauss, 1967) to keep track of emerging themes and links within the data.

Upon completion of the data gathering phase the primary investigator listened to all of the recordings of the data and supplemented her notes and memos as necessary. These analytic memos and notes were combined to create theoretical memos based on Glaser's (1998) idea that themes that emerge from data analysis can build theory. To accomplish this, memos were created for each emerging theme, with relevant supporting evidence added from the analytic memos. This enabled us to see which theories and concepts were most robust. Theories were identified based on the previous literature. For example, to identify whether theory of mind was present, the primary investigator analyzed the transcriptions for instances of the participants reporting perceived deception, motive, and intent of the IVR technology.

4. Findings

Participants reported that they contact client services for a variety of reasons, anything from making dinner reservations to concerns over bank or internet fraud; however, they did note that in the vast majority of situations they are emotionally invested in the outcome. People spoke about times when they had no power/heat during a winter storm, losing internet connection during exam time, and contacting companies about concerns about fraudulent charges or identity theft. In all of these cases the experience, regardless of the client service they receive, is fraught and stressful. Unsurprisingly then, people valued a speedy resolution above all else in their client service experiences. In these cases, clients are especially helpless; they cannot cancel their bank card themselves or fix the electricity or Wi-Fi in their homes; they are at the mercy of the companies.

Additionally, people who spoke about less dramatic issues, calling about a defective product or cancelling subscriptions, still reported feeling stress about their position of relatively limited power in the interaction. As

one participant in his 30s reported, "frustration mounts as time ticks away" without a resolution. Even for mundane activities such as cancelling a newspaper subscription, which two participants referenced, the frustration stems from the fact that they cannot handle the situation on their own and feel like they are beholden to the client service agent or IVR for resolution.

Because of their feelings of urgency, emotional investment, and perceived helplessness in the client service experience, contact with a live client service agent is the most sought after channel of client service communication, with many referring to it as "instinct" or "habit" to call to speak to a person. While it is the most sought after, the IVR feature that usually accompanies calls to companies received the most complaints from participants. People talked about long waits on hold and challenges in getting directed correctly by the IVR, but reported that they are willing to deal with this because they hold onto a belief that the live representative will help them. As one woman in her mid-20s said, "I feel like I can convince a person to help me". Two reported that they just prefer to complain to a live person. A young man in his 20s said that if he doesn't "have a person it feels like you are really on your own". Another female in her 20s said that she prefers to deal with a person instead of an IVR robot because "the robot has no sense of urgency. The robot doesn't care about me", which is an interesting comment to consider in light of the literature on human subconscious social responses to computers (Nass et al., 1996; Nass & Brave, 2005; Reeves & Nass, 1996). In this case, it would seem that the IVR robot has not been personified in a humanistic way, as she assumes it is less concerned or empathetic than a human would be. This of course does not imply that she did not respond socially to the IVR during the call, but rather to suggest that she sees a difference in her likelihood of success between the IVR and the live agent interactions. There is inherently more trust that the live representative will "see it through" as one respondent put it, compared to the IVR.

In addition to placing more trust in the human interactions, participants also noted that they are more likely to praise or think highly of a client service experience with a person as opposed to an IVR. The millennials in the tech savvy focus group all agreed that they will only "rave" about client service from a live agent because "I don't think anything of it if it's a robot". In other words, the IVR robot accomplishing client service seemed to be nothing special. They also noted that you can only get perks (e.g., free shipping, discounts) by phone because the agents are "invested in satisfying people right then and there". While this group agreed that the IVR can provide the efficiency that they seek in client service interactions, they noted that a robot technology will not garner "rave" reviews from them that a person would. Based on this it would seem that this group of millennials will discount robot technology to some degree. Again, the robot interaction is expected, or perhaps even unnoticed, in a way that interactions with live agents are not.

People see the live agent call as the final arbiter of their client service fate. As such, the IVRs that initiate these calls to a live agent were described as barriers to this service, a "roadblock" to getting help as one participant described it. Others described the time with IVRs as "feel[ing] like I'm being managed". A woman in her 30s said that the IVR was just "dragging the conversation out...you're just in line waiting", and she went on to say that the IVR does not actually lead to faster resolutions. It is clear from participants' language that they do not believe that the IVR will resolve their problems, and as such, they are things that must be endured to get through to the live agent.

And enduring them is hard work. People reported long wait times, questions that were irrelevant to their specific needs, and being "stuck in the loop of pressing buttons". In addition, the IVRs can be frustrating because of their long, but largely unhelpful, menu options. A woman in her 50s said that when there are too many menu options given verbally by the IVR, it is hard to remember and organize her options. In the third focus group (30–40 year olds) many noted that none of the menu options are relevant and "you can never remember which one" to press. These respondents said that they feel pressure in these situations because if you press the wrong button you will lose your place and probably have to restart the call. No one had the sense that they had the power to rectify a mistake in their IVR selections, which only further exacerbates levels of stress and frustration.

People also spoke about their frustration with company specific "keywords" or the "language" that IVRs employ. When the IVRs ask people to explain/state their problems, people were happy to have the chance to articulate their problems instead of pressing buttons, however, because of the robotic nature of the exchange many in the interviews talked about the importance of understanding the "language" or "keywords" of the IVR, which they noted can be hard to figure out. All focus group participants agreed with this as well. One woman in her 30s reported that she was frustrated because "I can't just tell my story" to the IVR in an open-ended way. And people expressed a disinclination to learn; as one woman in her 40s reported, "I don't want to have to know the language of the bank. I want the bank to know my language". This language is particularly hard to learn because of the infrequency of client service contact; one woman in her late 30s said that "I don't call these things enough to have their language be my language". Connected to this notion of language, some people noted that the IVRs can be challenging to deal with because they require the client to think about how to phrase the problem or question before speaking. One man in his 40s reported that he finds himself "verbally farting" when he tries to verbalize his problem to the IVR. Others reported the stress they felt when they do not really know the problem (i.e. trying to verbally explain a pipe or wiring problem in the house). These are real chal-

lenges for IVRs, which have a hard time following contradictions or stops/starts in language with accuracy.

This then introduces another layer to the power dynamic as the client must negotiate an IVR gatekeeper that is confusing to engage with because of long menus and unique keywords and language. Many respondents felt that the confusion, coupled with the long wait times, feels intentional; one referred to it as "strategic waiting" and others noted that it feels like "they don't want you to get through" to a live agent. We found that to manage this people have developed strategies for negotiating their IVR contacts that they alter depending on their experiences with the system in place. Thus, we see a difference between initial engagement with a company's IVR and the strategies they begin to employ as "time ticks away" and/or they become familiar with the IVRs language.

4.1. Initial Contact: On Guard but Polite

With the exception of a few people who reported that they never even listen to a IVR robot, opting out as soon as possible, most people reported that they will "entertain" the robot when they first hear it. Their reaction to their initial interaction with the robot seems to be to employ their social skills and work with the IVR to accomplish their goal. People used language that suggested that they are accommodating the machine, saying that they will "try it out", "try to listen", or "tolerate" the IVR when they first begin their client service experience. In these moments they will try to follow social norms and "engage" with the IVR. However, this moment of polite interaction with the IVR is short lived. Several people talked about how they had a tolerance for listening to only a certain number of prompts. One man in his 30s noted that "if I'm not at the right place in 2 or 3 beats, I'm not willing to keep trying things on a menu". Harkening back to the need for a quick resolution to their client service issue, people are only willing to interact politely with the IVR software for a short amount of time. However, again they cannot be rid of the IVRs altogether as they still need the robot to connect them to the live agent.

4.2. Continued Contact: Strategic Outmaneuvering

Once people have exhausted their patience with the IVRs they employ new strategies to manage the roadblock. Some reported that they will immediately start saying "agent", "representative", or pressing 0 in an effort to try to bypass the IVR robot. One woman in her 30s reported that she pretends "not to speak English or I say I have a rotary phone" so that she gets directed to the live agent. Interestingly in this phase they are not "tolerating" or "engaging" the IVR, but have instead moved to yelling at the robot, interrupting it, and/or pressing random buttons.

These experiences of increasing frustration and the strategic work-arounds they employ are remembered by these participants and they in turn inform their future client service interactions. With the exception of a few of the older participants who said that they always call client service out of "habit", most other people were willing to use different platforms to get to their end goal of a personalized, fast response. In fact, most participants were strategic in the channels they selected based upon 1) their prior experiences with the company's client service (i.e. knowing that the phone is notoriously slow) and 2) the nature of their question (ambiguous or multistep vs. straightforward; urgent vs. less urgent). In cases that they perceive to be ambiguous, involve a multistep process, and/or are urgent they prefer to call, while they may opt for email or live chat with straightforward questions. All participants reported that they consider these questions before deciding upon the channel to utilize. However it is important to note that a minority of participants noted that the technology channel they used is often preselected for them by the company (i.e. the company only offers an email or a web form). Thus the company can dictate the channels.

In most cases though, people are open to exploring other platforms as a way to optimize their results. First, every participant reported that he/she initiates any customer service contact by visiting the company website. Most navigated through the website a bit first, either to look for the customer service phone number (particularly popular with older respondents), or to search through the FAQs for an answer to their problem in an effort to avoid customer service altogether.

Others reported collective strategies that were beneficial both in terms of resolving their customer service issue as well as providing a psychological benefit. Some clients (40s and under) reported Googling their problem or question or looking on Twitter or Reddit to see how others are talking about this issue online. They liked doing this because it may help to lead them to a resolution. However, even if the search did not yield an answer, they reported that the results are still helpful because it puts the situation in perspective for them. The focus group of 30–40 year olds agreed with this, and as one male interviewee reported, it is helpful to know that others have had the problem too because it gets him out of thinking "why do these things always happen to me?" The focus group of tech savvy individuals picked up on this theme, with many noting that they look on Reddit to see that others are having the same problem and that it "keeps my emotions in check" to know that "others out there" have the same problem.

In general, what we see is participants' willingness to engage with the IVRs shifts as people gain more knowledge and understanding of how these processes will play out. While they may be willing to engage the IVR at the beginning, the perception of IVRs as gatekeepers means they have to create a work around to the system that is set up. They do this in a variety of ways, but most often either try to bypass the robot in the system or try another platform. Again, these strategies are devised solely to get a fast resolution to their often stressful client ser-

vice problems. As such, they make a calculated decision based on the importance of their problem and the ease of explanation. The more serious and confusing problems lead most people to call for a live agent, because the perception is that the live agent is most capable of solving the problem. Or at the least, the call will not end without some sort of response and understanding. For many people, just knowing that someone was working on a problem was enough to make them feel okay, but the perception is that only a person can create this feeling, not a robot.

4.3. Programmed Politeness

Many people in the interviews and focus groups made reference to the fact that occasionally they will encounter an IVR that engages in a programmed politeness, i.e. in its stilted robot voice the IVR apologizes to the client or expresses sympathy for the problem. This experience proved to have mixed effects with participants. The tech savvy focus group spent a great deal of time discussing this issue and largely agreed that it was possible that a robot could sympathize, or that someone could "program in sympathetic words" to make the overall experience more positive. In contrast several interviewees and participants in the focus group with 30–40 year olds said that they dislike the "fake empathy" of the robot. In both of these focus groups there was real debate about whether it mattered if the IVR was sympathetic as long as the problem was resolved, but there was no definitive answer on this as some felt that the resolution is the only thing that matters while others said that good customer service could improve any call.

We would like to acknowledge the limitations within the study. This was a small, exploratory study consisting of three focus groups and fifty interviews with time constraints. We have no claims to universality, as we used a non-random convenience sample and snowball method for recruitment. However, the data we gathered points very clearly to generalizability and it seems extensible. We hope to examine this further in future studies.

5. Discussion and Conclusion

This study explored how users manage their position communicating with IVR client service interfaces, with a focus on users' perceived amount of power. Our work first builds off of Nass et al.'s theories (1996, 2000, 2005) and those of Halpern and Katz (2013) and Katz and Halpern (2013) concerning social interactions with computers. Nass et al.'s work (1996, 2000, 2005) revolves around the media equation theory, the idea that people treat and behave toward computers the same way that they would another human. Research by Halpern and Katz (2013) and Katz and Halpern (2013) examines how users perceive robots and human-robot interaction. Similar to Nass et al., Katz and Halpern (2013) found that humans will ascribe humanlike qualities to robots even if they do not have a human-like appearance. However, the researchers also found that a user's willingness and likelihood to accept interaction with a robot is not dependent on the robot's human-likeness (Halpern & Katz, 2013).

Our findings echo aspects of these ideas of projecting human-like qualities onto IVR robots, specifically within initial interactions. H1 predicted that users interacting with IVRs for client service help will follow social norms and engage in polite interactions. We found this hypothesis to be partially supported, dependent on the users' experience with the IVR technology. What we saw from our studies is that, in novel IVR interactions, people revert to their social skills and human routines. At first, users are polite and willing to "engage" with the robot. Respondents indicated they would "try to listen" or "test out" the IVR service offering, which suggests that they are accommodating the system as if it were a live agent. In these cases, the users have limited power because the IVR is the gatekeeper and the situation is novel. People expressed frustration when engaging with these IVR technologies for the first time because the systems often required the understanding of "keywords" or the "language" of the automated system. Numerous study participants indicated that the infrequency of their customer service contact made it hard to learn the IVR's language and keywords. In these scenarios, there is a power imbalance between new IVR users and the IVR technology. Ultimately, to Castells' (2009) point, the protocols dictate the communication. As such, users rely on politeness and cooperation as a negotiating skill; they will listen through the interactions because they lack the power, in this sense the knowledge of an IVR system, to speed through the IVR prompts. This also provides insight into RQ1, which asked whether a users' behavior toward IVRs would change based on their perceived amount of power.

Through more frequent interactions, however, and as the conditions change and the IVR experience is no longer novel, people begin to see the technology only as a tool (i.e. less personal), specifically a tool that can be managed. Study participants indicated that, once they familiarized themselves with the IVR-friendly language and knew which buttons to press sequentially, IVR became less "stressful" and, in turn, "easier". Users stated they felt in "control" as they were able to easily go through the prompts and skip over irrelevant options after becoming familiar with the technology. With this reframing of the IVR as a merely a tool, people devise ways to work around the gatekeeper and begin to take control of the situation. They do this by dodging the IVR protocols by learning shortcuts they can apply to all IVR technology. These include pressing 0, yelling for an agent, or by turning the experience into a collective action whereby they search for human help online to bypass customer service altogether. By engaging in these behaviors, the clients develop more negotiating power as they no longer entirely unequal with the IVR technology.

With regards to the idea of IVR devices exhibiting emotion, empathy from computers was reported as

seeming "fake" and laden with ulterior motives. This perception of the technology's insincere empathy becomes particularly detestable in this second phase as it feels like a power play from the company. The idea is that this "programmed politeness" is a tool used by the companies to limit anger and discourage future action from users. The politeness is programmed to affect a certain outcome in the negotiation and it feels abhorrent to clients because it is false empathy, not honest or benevolent in intention. These ideas support H2, and users do employ theory of mind when interacting with IVR technology. This notion of the IVR being manipulative and disingenuous ties in with aforementioned research on cognitive science, specifically the theory of mind, which examines how people assign varying qualities and mental states to others (Kramar et al., 2012; Premack & Woodruff, 1978; Rilling et al., 2004; Thomas, 2001). The client's theory of mind encompassing IVR technology states that the robot is disingenuous and manipulative. Clients perceive the IVR's empathy as a way to deter future interactions and limit potential frustration: a company's deceptive manipulation tactic. From the user's perspective, once an IVR technology becomes a tool, it really cannot show true empathy.

Overall the customer service IVR experience is an interesting example of the ways in which people approach networked communication structures when they are powerless. What we found is that people are resourceful and able to develop means to enhance their position in the IVR negotiations. The interesting catalyst to these strategies is the reframing of the IVRs as a novel technology to an inhuman tool that must be circumvented. Our study also began to shed light on how media equation theory and theory of mind can be applied to human-robot interaction and perceptions of IVR technology. These vectors of social interaction analysis can be fruitfully expanded with an eye towards not only more fully understanding the social setting of an increasingly technologically mediated world but also to inform more humane usable interfaces to serve people more effectively.

Acknowledgments

The authors would like to thank Interactions LLC for sponsoring this research.

Conflict of Interests

The authors declare no conflict of interests.

References

Castells, M. (2009). *Communication power*. Oxford: Oxford University Press.

Center for Research on the Information Society. (2015). *Modalities of customer service* (CRIS Report). Morristown, NJ: CRIS.

Covan, E. K. (2007. The discovery of grounded theory in practice: The legacy of multiple mentors. In A. Bryant & K. Charmaz (Eds.), *SAGE handbook of grounded theory* (pp. 58–79). Thousand Oaks, CA: Sage Publications. doi:10.4135/9781848607941.n2

Fairhurst, G. T., & Chandler, T. A. (1989). Social structure in leader-member interaction. *Communication Monographs*, *56*(3), 215–239.

Foucault, M. (1984). *The history of sexuality* (Vol. 1). New York, NY: Pantheon Books.

Frith, C., & Frith, U. (2005). Quick guide: Theory of mind. *Current Biology*, *15*(17). doi:10.1016/j.cub.2005.08.041

Glaser, B. G. (1998). *Doing grounded theory: Issues and discussions*. Mill Valley, CA: Sociology Press.

Glaser, B. G., & Strauss, A. L. (1967). *The discovery of grounded theory: Strategies for qualitative research*. Chicago, IL: Aldine Publishing Company.

Goffman, E. (1959). *The presentation of self in everyday life*. New York, NY: Anchor Books, Doubleday.

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, *78*(6), 1360–1380.

Groom, V., Srinivasan, V., Bethel, C., Murphy, R., Dole, L., & Nass, C. (2011). Responses to robot social roles and social role framing. *International Conference on Collaboration, Technologies, and Systems*. doi:10.1109/CTS.2011.5928687

Halpern, D., & Katz, J. E. (2013). Close but not stuck: Understanding social distance in human-robot interaction through a computer mediation approach. *Intervalle*, *1*(3), 17–34.

Katz, J. E., Aspden, P., & Reich, W. (1997). Public attitudes toward voice-based electronic messaging technologies. *Behaviour & Information Technology*, *16*(2), 17–39. doi:10.1080/014492997119860

Katz, J. E., & Halpern, D. (2013). Attitudes toward robots suitability for various jobs as affected robot appearance. *Behaviour & Information Technology*. doi:10.1080/0144929X.2013.783115

Kitzinger, J. (1995). Introducing focus groups. *British Medical Journal*, *311*(1), 299–302. doi:10.1136/bmj.311.7000.299

Kramar, N., von der Putten, A., & Eimler, S. (2012). Human-agent and human-robot interaction theory: Similarities to and differences from human-human interaction. In M. Zacarias & J. Valente de Oliveira (Eds.), *Human-computer interaction: The agency perspective* (pp. 215–240). Berlin: Springer-Verlag.

Kraut, R., Patterson, M., Lundmark, V., Kiesler, S., Mukopadhyay, T., Scherlis, W. (1998). Internet paradox: A social technology that reduces social involvement and psychological well-being? *American Psychologist*, *53*(9), 1017–1031. doi:10.1037/0003-066X.53.9.1017

Lyons, M., Caldwell, T., & Shultz, S. (2010). Mind-reading and manipulation-is Machiavellianism related to theory of mind? *Journal of Evolutionary Psychology*, *8*(3), 261–274. doi:10.1556/JEP.8.2010.3.7

Nass, C., & Brave, S. (2005). *Wired for speech: How voice activates and advances the human computer relationship*. Cambridge, MA: MIT Press.

Nass, C., Fogg, B. J., & Moon, Y. (1996). Can computers be teammates? *International Journal of Human-Computer Studies*, *45*(6), 669–678. doi:10.1006/ijhc.1996.0073

Nass, C., & Moon, Y. (2000). Machines and mindlessness: Social responses to computers. *Journal of Social Issues*, *56*(1), 81–103. doi:10.1111/0022-4537.00153

Neuman, W. L. (2006). *Social research methods: Qualitative and quantitative approaches* (6th ed.). Boston, MA: Pearson, Allyn and Bacon.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *4*(1), 515–629. doi:10.1017/S0140525X00076512

Reeves, B., & Nass, C. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Stanford, CA: Center for the Study of Language Information and Cambridge University Press.

Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage*, *22*(4), 1694–1703. doi:10.1016/j.neuroimage.2004.04.015

Thomas, J. C. (2001). Collaborative innovation tools. In T. Terano, T. Nishida, A. Namatame, S. Tsumoto, Y. Ohsawa, T. Washio (Eds.), *New frontiers in artificial intelligence: Joint JSAI 2001 workshop post-proceedings* (pp. 27–34). Shimane: Springer.

Tiggemann, M., Gardiner, M., & Slater, A. (2000). "I would rather be a size 10 than get straight A's": A focus group study of adolescent girls' wish to be thinner. *Journal of Adolescence*, *23*(6), 645–659.

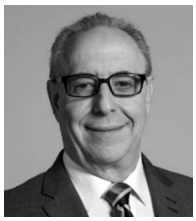
Tjosvold, D. (1984). Cooperation theory and organizations. *Impact Factor*, *37*, 743–767. doi:10.1006/jado.2000.0350

van Dijk, J. A. G. M. (2010). Review of "Manuel Castells, Communication Power". *Communications*, *35*, 485–489.

About the Authors

Jill Walsh's research revolves around media usage, particularly social media. She obtained a PhD in Sociology from Boston University in 2014, earlier earned a Master's in Public Policy from Brown University, and a BA from Harvard University. Prior to completing her graduate coursework, she taught, coached, and mentored high school students at the Noble and Greenough School for five years.

Brittany Andersen is a PhD student in the Emerging Media Studies Program (EMS) at Boston University College of Communication. At BU, her research focuses on user experience, mainly how users perceive, evaluate, and respond to different types of media and technology. She received her Master's from the EMS program in 2015, and has experience working on industry research and analyst relations teams.



James E. Katz, PhD., Dr.H.C., is the Feld Family Professor of Emerging Media at Boston University's College of Communication where he directs its Center for Mobile Communication Studies and Division of Emerging Media. He also serves as Distinguished Professor at the School of New Media School at Peking University. His research on social media and mobile communication has been internationally recognized and translated into many other languages. Among his books are *Machines That Become Us* and the *Handbook of Mobile Communication*.

Jacob Groshek is an Associate Professor of Emerging Media Studies at Boston University. His areas of expertise concern online and mobile media technologies as their use may relate to socio-political and behavioral health change at the macro and micro levels, and also include analyses of media content and user influence in social media. He uses a blend of interpretive as well as relatively advanced statistical tools for network analysis, forecasting, and explaining where and how the use of media has shaped the course of social, political, and health activity.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2018, Volume 6, Issue 3, Pages 69–72
DOI: 10.17645/mac.v6i3.1560

Commentary

The Implications of the FCC's Net Neutrality Repeal

Florian Schaub

School of Information, University of Michigan, Ann Arbor, MI 48109, US; E-Mail: fschaub@umich.edu

Submitted: 10 May 2018 | Accepted: 24 May 2018 | Published: 26 September 2018

Abstract
In December 2017, the Federal Communications Commission (FCC) repealed US net neutrality regulation. The author discusses the meaning and importance of net neutrality, the FCC's prior net neutrality rules and the implications of their repeal.

Keywords
Internet; net neutrality; public policy; regulation

Issue
This commentary is part of a multidisciplinary issue of *Media and Communication*, edited by Epp Lauk (University of Jyväskylä, Finland) and Raul Reis (Emerson College, USA).

1. Introduction

On December 14, 2017, the commissioners of the Federal Communications Commission (FCC)—the US regulatory agency for communications—voted to repeal the FCC's 2015 net neutrality regulation. The FCC's net neutrality repeal, as well as its prior regulations, were subject of heated public debate. So, what is net neutrality? What was repealed by the FCC? What are the implications of the FCC's net neutrality repeal?

2. What is Net Neutrality?

At its core, net neutrality is a principle requiring that the technology and entities that provide Internet connectivity and access should be impartial, or neutral, with respect to the communication and content for which that Internet access is used. This means that Internet service providers (ISPs) must treat all data and information on the Internet the same way (Krämer, Wiewiorra, & Weinhardt, 2013). They cannot discriminate or charge differently for different content or depending on who is providing a service.

By clearly distinguishing between the provisioning of Internet connectivity and Internet content, net neutrality ensures equal access to online content regardless of who is providing or requesting information. Net neutrality prevents ISPs from abusing their powerful position as mediator between consumers and content providers. This provides benefits for both consumers and online companies: consumers pay their ISP for Internet access and regardless of which ISP they choose, they will have the same level of access to online services, websites and apps. Online companies benefit because net neutrality creates a level playing field among small and large competitors. A startup with a good idea can create a website or app and be confident that anyone will have the same opportunity to access their service as those of the big established online companies. Non-profit organizations, individuals, and content creators can make their content available to the world without having to worry who will or won't be able to access it. Net neutrality ensures the openness and fairness of the Internet as well as facilitates innovation online. Many countries have created net neutrality laws or regulation to protect these tenets of the Internet, notably both the United States and Europe adopted net neutrality regulation in 2015.

3. Net Neutrality and the FCC

Not long after adopting net neutrality regulation for the United States in 2015, the FCC commissioners under FCC Chairman Ajit Pai voted in December 2017 to repeal the FCC's net neutrality regulation. This back and forth comes down to the question: Should ISPs be classified as common carriers or information services by the FCC?

3.1. The 2015 Net Neutrality Rules

In 2015, the FCC reclassified broadband ISPs as a telecommunications service ("common carrier") under Title II of the Communications Act of 1934 and Section 706 of the Telecommunications Act of 1996 to establish broader regulatory power over ISPs. This was a reaction to a court ruling that struck down the FCC's previous net neutrality principles, because the FCC had until then classified ISPs as "information services" under Title I of the Communications Act, which according to the court did not provide the FCC with the authority to enact net neutrality regulation for ISPs (*Verizon v. FCC*). Those earlier rules were put into place after some U.S. ISPs started throttling video streaming services, such as YouTube and Netflix, unless they entered into paid agreements with the ISPs. ISPs were effectively charging twice for the same service: consumers paid for access to the Internet and content providers had to pay to ensure their content would get to the consumers.

The 2015 FCC net neutrality regulation, titled the FCC's Open Internet Order (Federal Communications Commission, 2015), reclassified ISPs as telecommunication services—public utility providers—and prohibited ISPs from blocking access to legal content; throttling or otherwise impairing lawful Internet traffic; and establishing "fast lanes" for paid prioritization of some Internet traffic over other. This last point, however, did not mean ISPs could not engage in reasonable network management, but it prevented ISPs from pressuring content providers to enter into paid prioritization agreements. The order further stipulated that ISPs may not "unreasonably interfere with or unreasonably disadvantage" consumers' selection, access and use of online information, or the companies which provide those services (also called edge providers). ISPs further needed to disclose their network management and other practices.

3.2. The 2017 Net Neutrality Repeal

What changed with the FCC's 2017 vote? By adopting an order called "Restoring Internet Freedom", the FCC repealed the 2015 net neutrality rules in favor of returning to a "light-touch regulatory scheme" of the Internet (Federal Communications Commission, 2018a). More specifically, the FCC reversed the classification of ISPs as telecommunications services (common carriers), thus giving up its Title II authority over ISPs. As a result, ISPs are no longer regulated as public utilities and are now free to block or throttle Internet traffic, as well as offer paid prioritization, as long as they disclose their network management practices. This order also fully moves responsibility for investigating unfair trade practices by ISPs to the Federal Trade Commission (FTC).

FCC Chairman Pai (2018) argued that this change was necessary, because the 2015 net neutrality rules placed a high burden on companies and prevented investments in communications infrastructure and innovation. The FCC (2018a) and Pai (2018) claim that the Internet worked fine and prospered without the net neutrality rules before 2015, therefore no net neutrality regulation is required. Furthermore, they claim that the net neutrality repeal will help spur competition between ISPs and that competition will ensure fair and affordable access to the Internet across the United States.

4. Implications of the FCC Net Neutrality Repeal

Critics of the FCC's net neutrality repeal, myself included, do not share the FCC's optimistic outlook. The repeal of net neutrality heavily favors ISPs, with no benefits for consumers or even online companies. The FCC's net neutrality repeal places a lot of trust in ISPs to behave in consumer-friendly ways, which will likely amount to wishful thinking given ISPs track record in the United States. I expect that the implications of the FCC's net neutrality repeal will manifest through small and creeping changes rather than sudden shifts.

4.1. Higher (Indirect) Costs

The notion that the Internet prospered without much regulation and therefore does not need to be regulated going forward ignores the reality that ISPs in the United States have been increasingly exerting and abusing their powerful position as mediators between content (or edge) providers and consumers. It is very likely that ISPs will return to their practices of throttling high bandwidth services, such as video streaming, music streaming or video chatting, in order to pressure specific content providers into signing paid prioritization agreements to ensure access to the ISPs' networks and customers. Thus, without net neutrality ISPs can charge twice for the same service: their subscribers for Internet access and content providers for making sure that their data actually reaches the ISP's subscribers. While this may seem reasonable at first glance—after all those streaming services are bandwidth-intensive—an analogy reveals the absurdity of this situation. Imagine your power company would not just charge you for the electricity you use but in addition also require the manufacturer of your television, washing machine or electric car to pay any time one of their devices is used or charged, with the justification that watching TV or charging an electric car increases demand for electricity. While this is true, the customer is already paying for the use of this electricity! Similarly, Internet subscribers already have a paid agreement with an ISP that guarantees them a certain bandwidth and possibly transfer volume. ISPs additionally charging content providers for access to their customers constitutes "double dipping."

As a result, consumers may eventually end up paying more for online content, such as their Netflix subscriptions, because content providers would have to enter into separate agreements with different ISPs in the United States.

While large online companies may be able to shoulder such costs and make advantageous deals with ISPs, smaller companies, startups, and non-profit organizations may suddenly be faced with high costs to ensure that their services make it to consumers. Thus, rather than fostering innovation, the net neutrality repeal may stifle innovation in the online space by making it much more expensive with little benefit to consumers.

4.2. Bundled Offerings and Limited Internet Access

While double dipping is an issue, an even more concerning implication of the net neutrality repeal is that ISPs now have almost free reign over what content their subscribers will have access to, for instance due to how ISPs offer and package Internet access. With so called zero-rating offers, ISPs already exempt the use of certain online services from data caps.

Going forward ISPs may impose further requirements and charges for their zero-rating partners. The large U.S. ISPs have also become content providers through acquisitions of NBC Universal (by Comcast), Time Warner (by AT&T) and other content providers. Thus, ISPs have strong incentives to give preferable treatment to their own content, as well as enter into mutual agreements with other entities that also control access to Internet subscribers or content subscribers.

Zero-rating may lead towards even more tailored and tiered bundled Internet access offerings. In the future, some online services may come for free with your basic Internet package due to respective agreements. If you want access to that other video platform you will have to get the extended package. You want access to the whole Internet? Sure, that's still available—just sign up for the premium package. In such a model, ISPs would effectively replicate the antiquated television channel model for Internet access.

Based on tiered pricing, ISPs could curate what online content and services most people will have access to, and which ones will only be available to those who are willing to and can afford to pay extra. People of low socioeconomic status may be especially disadvantaged by having less access to the full internet. While tiered access models may exacerbate socioeconomic differences, ISPs could also suppress certain websites and online services, as long as they transparently disclose those practices in their terms of service. This would have a detriment effect for equal access to online resources, with serious implications for access to knowledge, free speech, and democratic participation.

4.3. Little Competition among ISPs

While increased competition among ISPs in the United States would be welcome, there is little indication that the net neutrality repeal would substantially spur competition among ISPs or incentivize ISPs to substantially invest in improving their network infrastructure. In most parts of the United States, especially in rural areas, people have little to no choice regarding ISPs. According to a recent FCC report (Federal Communications Commission, 2018b), in 2016, 63.2% of developed census blocks in the United States had only one cable provider, 4.1% had two or three, 32.7% had none. The report shows similarly sparse competition for other Internet access technologies.

This lack of competition among ISPs means that transparency about network management practices alone is not sufficient to protect consumers. Most U.S. consumers will not be able to switch ISPs if they disagree with their practices or pricing model. Furthermore, even if they are so lucky to be able to choose among two or three ISPs, that choice may not be meaningful if those ISPs implement similar network management practices. As long as those practices are made transparent and are not clearly deceptive or unfair—the only cases in which the FTC can become active—ISPs will be able to get away with throttling and other practices violating net neutrality.

5. Conclusion

While the FCC's net neutrality repeal constitutes a major setback for a neutral and open Internet, it will likely not mean the end of net neutrality in the United States. At the time of writing, multiple lawsuits against the FCC's decision have been filed by consumer advocacy groups, attorneys general from multiple U.S. states, as well as technology companies. Some states, including Washington and Oregon, have passed state laws prohibiting ISPs from blocking or throttling Internet traffic. In other states, including California and New York, net neutrality state laws have been introduced. The governors of New Jersey, Montana and other states have signed executive orders requiring ISPs to adhere to net neutrality principles if they do business with the state. Net neutrality bills have also been proposed in congress. In May 2018, the U.S. Senate voted to overturn the FCC's Restoring Internet Freedom Order, but at the time of writing it is doubtful whether the same vote would reach quorum in the House of Representatives. The FCC's net neutrality repeal was misguided and will hopefully be rectified in the United States and not replicated in other countries. Net neutrality is essential to ensure open exchange, free speech, and innovation online by preventing the entities that provide access to the Internet from discriminating among traffic, information, content providers or individuals.

Conflict of Interests

The author declares no conflict of interests.

References

Federal Communications Commission. (2015). *Protecting and promoting the open Internet* (80 FR 19737). Washington, DC: US Government.

Federal Communications Commission (2018a). *Restoring Internet freedom* (83 FR 7852). Washington, DC: US Government.

Federal Communications Commission (2018b). *Internet access services: Status as of December 31, 2016*. Washington, DC: US Government.

Krämer, J., Wiewiorra, L., & Weinhardt, C. (2013). Net neutrality: A progress report. *Telecommunications Policy*, *37*(9), 794–813.

Pai, A. (2018). *Statement of Chairman Ajit Pai Re: Restoring Internet Freedom* (WC Docket No. 17–108). Washington, DC: Federal Communications Commission.

Verizon v. FCC 2014 11-1355 United States Court of Appeals for the District of Columbia Circuit.

About the Author

Florian Schaub is Assistant Professor of Information and Electrical Engineering and Computer Science at the University of Michigan. Dr. Schaub's research focuses on understanding and supporting people's privacy and security behavior and decision making in complex socio-technological systems. His research interests span privacy, human-computer interaction, and emergent technologies, such as the Internet of Things. Dr. Schaub received his doctoral degree in Computer Science from the University of Ulm, Germany.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2018, Volume 6, Issue 3, Pages 73–76
DOI: 10.17645/mac.v6i3.1596

Commentary

The New Frontier in Communication Research: Why We Should Study Social Robots

Jochen Peter * and Rinaldo Kühne

Amsterdam School of Communication Research (ASCoR), University of Amsterdam, 1018 WV Amsterdam, The Netherlands; E-Mails: j.peter@uva.nl (J.P.), r.j.kuhne@uva.nl (R.K.)

* Corresponding author

Submitted: 4 June 2018 | Accepted: 11 June 2018 | Published: 27 September 2018

Abstract
Social robots—robots that are made for interaction with humans—are becoming increasingly popular. In contrast to other disciplines, however, communication research has been slow in studying them. In our view, there are at least three theoretical reasons for communication researchers to deal with social robots. First, social robots challenge our notions of medium and media. Second, social robots challenge our understanding of the communication partner. Finally, social robots challenge our notions of the boundaries of communication. We therefore believe that social robots should play a more central role in communication research than it is currently the case.

Keywords
artificial intelligence; communication science; human-machine interaction; human-robot interaction; social robots

Issue
This commentary is part of the Multidisciplinary Issue of Media and Communication, edited by Epp Lauk (University of Jyväskylä, Finland) and Raul Reis (Emerson College, USA).

1. Introduction

Due to ground-breaking advancements in computing, sensor technology, and artificial intelligence (e.g., Eberl, 2016), robots are nowadays increasingly designed for interaction with human beings (for a recent review see, e.g., Broadbent, 2017). These social robots integrate multiple communication modalities (e.g., vision, speech, touch) and can, once properly programmed, relate to human beings in meaningful ways (Broadbent, 2017; Fong, Nourbakhsh, & Dautenhahn, 2003; Lee, Peng, Jin, & Yan, 2006). As a result, social robots are already used as language tutors, as well as companions for elderly people and children with autism spectrum disorder (Cabibihan, Javed, Ang, & Aljunied, 2013; de Graaf, Allouch, & Klamer, 2015; Han & Kim, 2009). Moreover, a growing number of interactive 'smart' toys, which typically rely on social-robot technology, are currently entering the market (Future of Privacy Forum & Family Online Safety Institute, 2016; Peter, Kühne, Barco Martelo, De Jong, & Van Straten, in press). Finally, scholars and public commentators expect that, in the future, people will progressively encounter social robots as companions, collaborators and colleagues (e.g., Dautenhahn, 2007; Mols & Vergunst, 2017). Against this backdrop, several observers consider social robots a key future technology (Barnatt, 2015; Ross, 2016).

While engineering sciences and robotics have been investigating social robots for some time, communication research's response to the emergence of social robots has tended to be rather slow and scattered, some notable exceptions notwithstanding (e.g., Sandry, 2015; Zhao, 2006). This is surprising because in any interaction between social robots and humans some type of communication is essential, regardless of whether this communication is verbal or non-verbal. Moreover, various scholars have recently called for more attention of communication researchers to intelligent machines (e.g., Gunkel,

2012; Jones, 2014). In a paper published more than ten years ago and aptly titled "Humanoid Social Robots as a Medium of Communication", Zhao (2006, p. 402), for example, observed: "[The] emerging movement of social roboticization is causing a fundamental change in the meaning of social interaction and the nature of human communication in society". Focusing more broadly on semi-intelligent machines and 'smart' devices, Gunkel (2012, p. 2) has requested that "[c]ommunication studies…must come to terms with this development and reorient its theoretical framework". Thus, we are not the first to link social robots to communication and to propose that communication researchers should focus on human-machine communication; one of us has also requested already elsewhere that communication researchers pay attention to social robots, notably in research on young people (Peter, 2017). Based on existing research (e.g., Gunkel, 2012; Guzman, in press; Zhao, 2006), we rather believe that at least three important reasons for communication researchers to study social robots need to be (re)emphasized, especially now that the developments surrounding social robots are becoming increasingly powerful and pervasive (for an elaboration of the first two points below, see also Peter, 2017).

2. Three Reasons to Study Social Robots

First, social robots challenge our notions of medium and media. As Zhao (2006, p. 402) has succinctly noted, "social robots…are not a medium through which humans interact, but rather a medium with which humans interact". Social robots thus do not just function as mere transmission channels—a conceptual and theoretical problem that has been described also for computers in particular and media in general (Cathcart & Gumpert, 1983; Gunkel, 2012). Rather, social robots transcend the role of a medium because they can be both senders and receivers and acquire the status of social actors (e.g., Gunkel, 2012; Guzman, in press). Empirical research within the computers-are-social-actors paradigm has solidly demonstrated that human beings treat computers, and media more generally, as social actors and eventually as if they were human (e.g., Reeves & Nass, 1996). Given the vastly expanded abilities and characteristics of social robots, the idea of a medium as a communication partner thus deserves more attention (e.g., Zhao, 2006).

Second, social robots challenge our understanding of the communication partner. The vast amount of communication research—be it on interpersonal, computer-mediated, or mass media communication—seems to assume (at least implicitly) that communication takes place between two or more human beings (e.g., Guzman, in press). A social robot, however, can be seen as "another kind of communicative Other—who confronts human users, calls to them, and requires an appropriate response", as Gunkel (2012, p. 21) put it, referring to computers more generally. The communication partner is thus no longer human. Accordingly, social robots force us to reconsider the notion that the communicative other is typically human (Gunkel, 2012). Attention to this major shift merges partly with what has been called the 'nonhuman turn' (Grusin, 2015a). The nonhuman turn currently takes place in various fields in the social sciences and humanities that are "engaged in decentering the human in favor of a turn toward and concern for the nonhuman" (Grusin, 2015b, p. vii). We certainly do not advocate abandoning the human in communication research. Similar to others (Gunkel, 2012; Guzman, in press; Zhao, 2006), however, we do believe that an extension of our theoretical and empirical research to non-human communication partners is not only timely but will also advance the field of communication research as a whole.

Third, social robots challenge our notions of the boundaries of communication. Social robots vary in their morphology from anthropomorphic (with human-like features), to zoomorphic (with animal features), to caricatured (similar to animation figures), to functional (with machine-like features) (Fong et al., 2003). As these robots may increasingly feature advanced and perhaps even unique skills, communication with them may go beyond what we currently know about human-human, human-animal, human-agent, or human-machine communication (Sandry, 2015). More specifically, according to some observers (e.g., van Bergen, 2016), social robots may in the future have better language and visual skills than human beings. With the advancement of social robots that are supposed to read human emotions and respond to them, such as Pepper (SoftBank Robotics), communicative possibilities may thus emerge that may exceed the boundaries of human communication (Sandry, 2015). This development may force us to confront, also in human-robot communication, what Sundar (2008) has termed the 'ftf fallacy' in reference to the relation between computer-mediated communication (CMC) and face-to-face (ftf) communication. According to Sundar (2008, p. 59), the ftf fallacy implies that "[f]tf is the gold standard, and all CMC innovations, situations, and devices are measured against this standard". Being aware of the ftf fallacy in human-robot communication means acknowledging that communication with social robots may be different from ftf communication and not necessarily comparable with it. Applied to human communication more generally, it means that human-robot communication finds its equal place next to human–human communication (e.g., Guzman, in press).

3. Conclusion

In summary, the emergence of social robots challenges three paradigmatic assumptions in communication research—about the medium, the communication partner and the boundaries of communication. We believe that it is crucial that communication research broadens its scope to the study of social robots in order to de-

velop more comprehensive communication theories (for this request more generally focused on human-machine communication see, e.g., Gunkel, 2012; Guzman, in press). However, we warn researchers also against an exaggerated tech optimism, notably against overestimating the ease of doing research with current social robots (Belpaeme et al., 2013), and falling into the trap of technological determinism. Still, we are convinced that social robots should receive more attention and be given a more central position in communication research.

Acknowledgements

Work on this commentary was made possible by funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme, under grant agreement no. 682733 to the first author.

Conflict of Interests

The authors declare no conflict of interests.

References

Barnatt, C. (2015). *The next big thing: From 3D printing to mining the moon*. Retrieved from www.explainingthefuture.com

Belpaeme, T., Baxter, P., de Greeff, J., Kennedy, J., Read, R., Looije, R., . . . Zelati, M. C. (2013). Child-robot interaction: Perspectives and challenges. In *Proceedings of the 5th International Conference on Social Robotics* (Vol. 8239, pp. 452–459). New York, NY: Springer.

Broadbent, E. (2017). Interactions with robots: The truths we reveal about ourselves. *Annual Review of Psychology*, *68*(1), 627–652. https://doi.org/10.1146/annurev-psych-010416-043958

Cabibihan, J.-J., Javed, H., Ang, M., & Aljunied, S. M. (2013). Why robots? A survey on the roles and benefits of social robots in the therapy of children with autism. *International Journal of Social Robotics*, *5*(4), 593–618. https://doi.org/10.1007/s12369-013-0202-2

Cathcart, R., & Gumpert, G. (1983). Mediated interpersonal communication: Toward a new typology. *Quarterly Journal of Speech*, *69*(3), 267–277. https://doi.org/10.1080/00335638309383654

Dautenhahn, K. (2007). Socially intelligent robots: Dimensions of human-robot interaction. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *362*(1480), 679–704. https://doi.org/10.1098/rstb.2006.2004

de Graaf, M. M. A., Allouch, S. B., & Klamer, T. (2015). Sharing a life with Harvey: Exploring the acceptance of and relationship-building with a social robot. *Computers in Human Behavior*, *43*, 1–14. https://doi.org/10.1016/j.chb.2014.10.030

Eberl, U. (2016). *Smarte Maschinen. Wie künstliche Intelligenz unser Leben verändert* [Smart machines. How artificial intelligence changes our lives]. Munich: Hanser.

Fong, T., Nourbakhsh, I., & Dautenhahn, K. (2003). A survey of socially interactive robots. *Robotics and Autonomous Systems*, *42*(3/4), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

Future of Privacy Forum, & Family Online Safety Institute. (2016). Kids and the connected home: Privacy in the age of connected dolls, talking dinosaurs, and battling robots. *Future of Privacy Forum*. Retrieved from https://fpf.org/wp-content/uploads/2016/11/Kids-The-Connected-Home-Privacy-in-the-Age-of-Connected-Dolls-Talking-Dinosaurs-and-Battling-Robots.pdf

Grusin, R. (Ed.). (2015a). *The nonhuman turn*. Minneapolis, MN: University of Minnesota Press.

Grusin, R. (2015b). Introduction. In R. Grusin (Ed.), *The nonhuman turn* (pp. vii–xxix). Minneapolis, MN: University of Minnesota Press.

Gunkel, D. J. (2012). Communication and artificial intelligence: Opportunities and challenges for the 21st century. *Communication +1*, *1*(2012), 1–25.

Guzman, A. L. (in press). What is human-machine communication, anyway? In A. L. Guzman (Ed.), *Human-machine communication: Rethinking communication, technology, and ourselves*. New York, NY: Peter Lang.

Han, J., & Kim, D. (2009). r-Learning services for elementary school students with a teaching assistant robot. In *Human-robot interaction (HRI)* (pp. 255–256). La Jolla, CA: IEEE.

Jones, S. (2014). People, things, memory and human-machine communication. *International Journal of Media & Cultural Politics*, *10*(3), 245–258. https://doi.org/10.1386/macp.10.3.245_1

Lee, K. M., Peng, W., Jin, S.-A., & Yan, C. (2006). Can robots manifest personality? An empirical test of personality recognition, social responses, and social presence in human-robot interaction. *Journal of Communication*, *56*(4), 754–772. https://doi.org/10.1111/j.1460-2466.2006.00318.x

Mols, B., & Vergunst, N. (2017). *Hallo robot. De machine als medemens* [Hello robot. The machine as a fellow human]. Amsterdam: Nieuw Amsterdam.

Peter, J. (2017). New communication technologies and young people: The case of social robots. In R. Kühne, S. E. Baumgartner, T. Koch, & M. Hofer (Eds.), *Youth and media. Current perspectives on media use and effects* (pp. 203–217). Baden-Baden: Nomos.

Peter, J., Kühne, R., Barco Martelo, A., De Jong, C., & Van Straten, C. L. (in press). Asking today the crucial questions of tomorrow: Social robots and the internet of toys. In G. Mascheroni & D. Holloway (Eds.), *The internet of toys: Practices, affordances and the political economy of children's play*. Basingstoke: Palgrave Macmillan.

Reeves, B., & Nass, C. (1996). *The media equation. How people treat computers, television, and new media*

like real people and places. Cambridge: Cambridge University Press.

Ross, A. (2016). *The industries of the future*. New York, NY: Simon & Schuster.

Sandry, E. (2015). *Robots and communication*. Basingstoke: Palgrave Macmillan.

Sundar, S. S. (2008). Self as source: Agency and customization in interactive media. In E. A. Konijn, S. Utz, M. Tanis, & S. B. Barnes (Eds.), *Mediated interpersonal communication* (pp. 58–74). New York, NY: Routledge.

van Bergen, W. (2016). *De robots komen eraan! Feit en fictie over de toekomst van intelligente machines* [The robots are coming! Fact and fiction about the future of intelligent machines]. Amsterdam: Business Contact.

Zhao, S. (2006). Humanoid social robots as a medium of communication. *New Media & Society*, *8*(3), 401–419. https://doi.org/10.1177/1461444806061951

About the Authors

Jochen Peter (PhD, University of Amsterdam, 2003) is a Full Professor in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. His research deals with young people's use of new communication technologies and its consequences for their psycho-social development.

Rinaldo Kühne (PhD, University of Zürich, 2015) is an Assistant Professor in the Amsterdam School of Communication Research (ASCoR) at the University of Amsterdam. His research focuses on media effects on young people, the communication between humans and artificial agents, cognitive and emotional mechanisms of media use and effects, and methods in communication research.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2018, Volume 6, Issue 3, Pages 77–82
DOI: 10.17645/mac.v6i3.1601

Commentary

News, Ads, Chats, and Property Rights over Algorithms

Jan Kleinnijenhuis

Department of Communication Science, Vrije Universiteit Amsterdam, 1081HV, Amsterdam, The Netherlands;
E-Mail: j.kleinnijenhuis@vu.nl

Submitted: 6 June 2018 | Accepted: 11 June 2018 | Published: 28 September 2018

Abstract
The success of tech firms rests on their ownership of the algorithms for operating new platforms for the interactions among five groups of stakeholders in the markets of news, ads, and chats: stakeholders from the spheres of politics, journalism, the citizenry, the tech firms themselves, and other firms. Recent regulations that touch on property rights such as the German Netzwerkdurchsetzungsgesetz and the European Directive on Copyright in the Digital Market have turned ownership of algorithms into exclusive ownership. Thereby tech firms obtain also the right to censor and the exclusive right to micro-target clients for advertisers. Coase's theorem is used to discuss alternative allocations of property rights that could improve the quality of news, ads, and chats.

Keywords
algorithms; networks; property rights; social media; tech industry

Issue
This commentary is part of the Multidisciplinary Issue of Media and Communication, edited by Epp Lauk (University of Jyväskylä, Finland) and Raul Reis (Emerson College, USA).

1. Introduction

"I'm a lawyer and I have no idea what that means" declared Republican Senator Lindsey Graham while holding up the Facebook Terms of Service in the Congressional Hearing regarding Facebook on April 10th, 2018 (C-Span, 2018). In spite of such objections, the US Congressional Hearing, the German Netzwerkdurchsetzungsgesetz (NetzDG)[1] of January 2018, the General Data Protection Regulation (GDPR)[2] of May 2018, and the Directive on Copyright in the Digital Market ("Copyright Directive", approved by the European Parliament in September 2018; expected implementation in laws of the European Union (EU) member states in 2019)[3], all push in the direction of giving even more rights to tech firms: first, the right to use their algorithms to censor, as well as the exclusive right to use their algorithms to micro-target their users with messages from firms and advocacy groups who pay for this.

NetzDG encourages internet platforms to censor. Fines of up to 5 million euros may be imposed (4.2) if they do not remove client postings with offending content. NetzDG may promote advocacy groups to accuse their political opponents of hate speech, so as to cause tech firms to block or remove opinionated messages of their opponents. In the Congressional hearing, Facebook's CEO, Mark Zuckerberg announced that, by the end of 2018, 20,000 people will work on content review. Ultimately more advanced AI algorithms "would block millions of fake accounts each day at the point of creation before they do any harm" (Facebook, 2018, p. 110). The Copyright Directive forces tech firms explicitly to develop "effective content recognition technologies" which enable censorship. These technologies should enable tech firms to let its users pay for linking to news items from media companies (11, "link tax") and to prevent its users from uploading or disseminating copyrighted materials (13, "upload filter"). According to a group of the inter-

[1] https://www.gesetze-im-internet.de/netzdg/NetzDG.pdf
[2] https://gdpr-info.eu/
[3] Copyright Directive proposal 2016: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52016PC0593. Amendments European Parliament September 2018 http://www.europarl.europa.eu/sides/getDoc.do?type=TA&language=EN&reference=P8-TA-2018-0337

net's architects, including Vincent Cerf (internet protocol TCP/IP) and Timothy Berners-Lee (World Wide Web protocol HTTP), this is "an unprecedented step towards the transformation of the internet from an open platform for sharing and innovation, into a tool for the automated surveillance and control of its users" (Cerf et al., 2018).

An exclusive ownership right given to tech firms to use their algorithms to micro-target users with third party messages swings the pendulum back. In 2013, Facebook, for example, offered its users the right to use Facebook's GraphSearch algorithm to explore the preferences of friends of friends. According to Alexandr Kogan, who used GraphSearch to scrape Facebook user data for CambridgeAnalytics, scraping was quite common at that time. The try-out was stopped in December 2014 because of privacy issues and because its use made Facebook slow. In 2018, Mark Zuckerberg could easily state, therefore:

> In 2014, to prevent abusive apps, we announced that we were changing the entire platform to dramatically limit the data apps could access. Most importantly, apps like Kogan's could no longer ask for data about a person's friends unless their friends had also authorized the app. (Zuckerberg, 2018)

Future advertisers and campaigners will have to pay the tech firms to use their graph search algorithms for micro-targeting. As a result, the tech firms earn the largest share of advertising revenues, although algorithm owners may decide to let successful vloggers share in advertisement revenues once they attract many new clients (YouTube), or put advanced social media monitoring tools such as CrowdTangle at the disposal of media companies so as to maximize the audience for their news items on their social media pages (Facebook). Thus, tech firms who own the search algorithms that link their users to news content provided by media companies help to increase the traffic to the websites of these media companies, which enables the latter to sell advertisement space on their own websites. In advance, media companies did contribute to the traffic towards the search algorithms of tech firms, and therefore to the advertising revenues of tech firms, with the cumulative supply of their latest news content. Only a fraction of the search algorithm users in search of news content click on the news snippets or news summaries from media companies that tech firms show to them. Only a fraction reaches the websites of media companies. This explains why the advertisement revenues of tech firms went up disproportionally as compared to those of media companies. This explains also why tech companies are opposed to the Copyright Directive that would reward media companies for the traffic that tech companies can attract by providing links to, or using links to, news content provided by media companies. According to UK Music, Google alone spent more than 30 million euro in its successful lobby to depict the Copyright Directive as the end of an open internet with free links to other websites (UK Music, 2018). Google exploited its ownership of the Pagerank algorithm by inducing the main German media companies, who feared the removal of online links to their news articles on NetzDG copyright grounds, to undermine NetzDG by entering into zero price licenses with Google. With $7.4 million in litigation costs against Google, the German association of media companies could use NetzDG to let Google pay out €0.7 million to media companies who were not satisfied with a zero price license (Ehle, 2018). The legitimacy of media companies in such judicial battles is low because news content is delivered by journalists, while media tycoon companies appear to want to become tech firms themselves by investing less in journalism than in online presence and marketing.

Here we will take Coase's theorem as the point of departure to answer the research question whether alternative allocations to stakeholders of property rights over algorithms would result in a lower price, or a higher quality of news, ads, and/or chats.

2. Coase Theorem

Ronald H. Coase (1910–2013) received the 1991 Noble prize "for his discovery and clarification of the significance of transaction costs and property rights for the institutional structure and functioning of the economy". Property rights apply to palpable objects that you are holding in your hands, but also to land, labor force, money, assets, debt certificates, futures, patents, radio frequencies, and network algorithms.

The Coase theorem holds that if property rights are allocated and protected by a Sovereign, and transaction costs are independent of who holds the property rights, then parties will negotiate about the price of benefitting from the use of such property. They will agree on the same price, regardless of who holds the property rights—what differs is merely who has to pay the price (Coase, 1960/2013). A classic "reciprocal" example is that river delta inhabitants who want to prevent the pollution of the river will be ready to pay the same price to mountain inhabitants who own the river, as the price that mountain inhabitants are ready to pay to dump their polluted waste in the river owned by the river delta inhabitants.

Coase developed his theorem in the context of his investigations into the lack of explicit ownership over radio frequencies in the first half of the previous century (Coase, 1959/2013). To be able to develop a judicial framework to settle disputes about the interference of different radio signals, the government should use its sovereignty to allocate property rights, just as Thomas Hobbes' Sovereign (Hobbes, 1660/2002). This could be achieved by having auctions to sell radio frequencies to the highest bidder. Without property rights, the legal system would keep producing inconsistent jurisdiction as to whether government agencies had any discretion to refuse a radio license, and whether refusing a radio li-

cense was a violation of the constitutional freedom of the press.

Just as radio programs relate to radio frequencies which allow the targeting of citizens and firms within a geographical area, so do chats, ads, news, and data relate to algorithms which allow for the micro-targeting of specific firms or citizens. The claim of this article is that, just as disputes over the constitutional freedom of the press were solved by government regulations to establish ownership of radio frequencies, disputes over the freedom of expression versus surveillance, censorship, and control can be solved in part by new government regulation regarding property rights over network algorithms.

The Coase theorem urges us to ask which stakeholders are involved, whether property rights are sufficiently clear and whether transaction costs sufficiently low to negotiate a deal.

3. Stakeholders to Whom Property Rights Could Be Allocated

Tech firms own network algorithms. Basically, a network algorithm relies on relationships R between concepts *X* and *Y*, that could be modeled as RDF-triplets. It derives "transitive" logical conclusions, for example whether *X* will presumably like *Z* based on information about the relationship of *X* to *Y*, and of *Y* to *Z*, e.g., conclusions about a book *Z* that *X* may want to read, given the resemblance between book *Z* and a book *Y* that was already bought by *X*. The network algorithm that underlies the right-hand side of Google pages, for example, often offers a better comparison of industries and products than web pages of individual organizations filled with PR clichés.

To think about possible alternative allocations of property rights over network algorithms in the market of news, advertisements, and chats, Figure 1 shows a simplified model with five groups of stakeholders from the spheres of the network algorithm firms themselves, other firms, politics, journalism, and the citizenry.

Within a few decades, network algorithm firms—e.g., Google, Amazon, Facebook, Yahoo Finance, Bloomberg, Thomson Reuters' Datastream, and RELX's LexisNexis—became "semi-benevolent information monopolies" (Karpf, 2018) which established new relations with the other four types of actors. More importantly, they also afforded new relations between the latter. To relate Figure 1 above to its explanation below, italics will be used for the newly afforded relationships between the stakeholders that are printed as arrow labels in Figure 1.

Tech firms offer the *platforms*, *data*, and *information*, on which firms, financial markets, citizens and politicians rely. They deliver citizens a new platform to chat with other citizens, sometimes in their roles as voters or consumers. They allow firms, both in the *B2B*- and

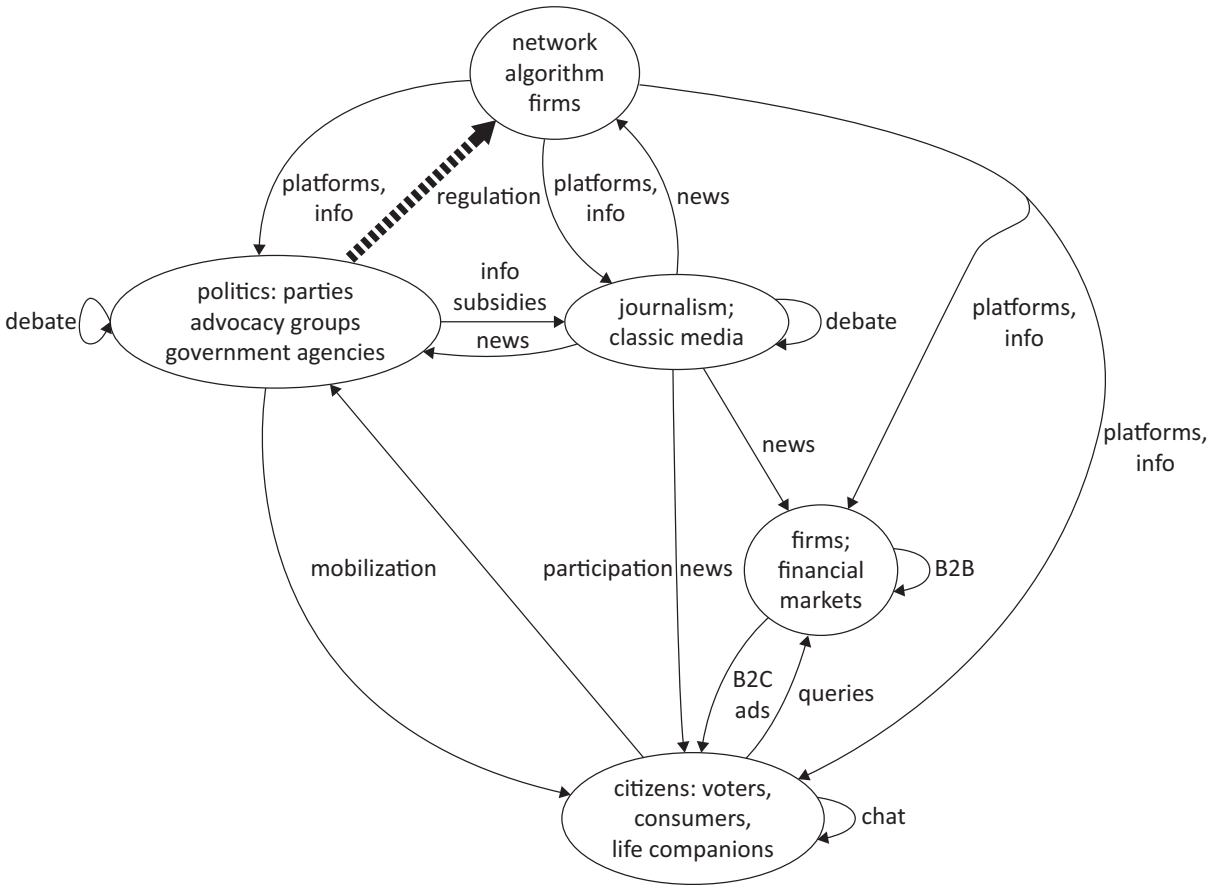


Figure 1. Reciprocal relations between network algorithm stakeholders in the market of news, ads, and chats.

B2C-markets, to micro-target consumers with *ads*. Their social media algorithms offer parties, advocacy groups, and government agencies a means to *mobilize* citizens to *participate*, share, contribute or vote. Tweets and Facebook posts from politicians have become a new political *information subsidy* (Cook, 1998) to journalism, in combination with the social media metrics to guess their impact. It is a weird symbiotic relationship, in which attack tweets especially, among others on media and journalists, increase likes and retweets (Lee & Xu, 2018). The resulting social media storm is often a basis for *news* in classic media, even when the tweets were not initially deemed newsworthy (Wells et al., 2016).

Social media have become a major portal to access *news* items from classic media, although online surveys somewhat overestimate their use (Newman, Fletcher, Kalogeropoulos, Levy, & Nielsen, 2018). The news supply of classic media is split into separate news items with a standard format whose clickbait is routinely assessed with A/B tests. Relatively slow *news* in classic media about political and financial networks nevertheless exerts not only an influence on public support for political candidates (Wells et al., 2016), but also a long-term influence on financial markets (Kleinnijenhuis, Schultz, Oegema, & van Atteveldt, 2013; Kräussl & Mirgorodskaya, 2017). The relevance and influence of news is not captured in price information in the way that modernists like Smith ("the invisible hand") and Karl Marx ("hinter dem Rücken der Produzenten") predicted. Classic media are still the major source of online news (Thurman & Schifferes, 2012), but their advertising revenues are decreasing because the advertising revenues of the tech firms who disseminate their news are increasing.

4. Price and Quality of News, Ads, and Chats After Alternative Allocations of Algorithm Ownership

The puzzle to solve is whether an alternative allocation of algorithm ownership would result in a lower price, or a higher quality, of news, ads, and chats.

A re-allocation that simply would not work is to split giant tech firms into parts. The parts would have a strong incentive to cooperate to serve advertisers with potential customers everywhere and to serve users with friends everywhere. Prices would remain largely unaltered, in line with the Coase theorem. A functional split, similar to the split in the banking sector between investment banks and consumer banks, would not work either, because there is no business model for news or chats without ads, and no model of democracy based on quality news without chats and debate.

The European GDPR grants internet users ownership over their personal data, without providing them with an efficient means to sell it. Therefore, they have to deliver them for free by clicking a thousand times that they accept all cookies before they can carry out business as usual.

More radical ways to split up property rights may not work either because transaction costs, which are primarily energy costs in the case of the internet, would be much higher. The enormous energy costs of the current internet are negligible when compared to an internet in which unsupervised high-frequency crypted transactions between all users would be required (i.e., blockchain technology) and in which moreover blocking or removing of specific content would be both content specific (e.g., censorship, copyright) and user specific (e.g., various ratings of social credit). For the same reason, incredibly high transaction costs would result if property rights of graph search algorithms were granted to firms, journalists, and political agencies so that individual internet users would have to pay time and again to avoid being micro-targeted. High transaction costs that threaten the internet will result also if EU member states would make a mess out of property rights by implementing the link tax and the upload filter from the Copyright Directive ambiguously or inconsistently so that many different parties in EU member states may go to court to claim on very different grounds that tech companies violated their property rights over news items, texts, music, films, video clips, art, patents, intellectual property, or art.

Re-allocating the ownership of a part of the network algorithms invented by tech firms to stakeholders from the spheres of journalism, politics, firms and the citizenry could however prevent that tech firms obtain a right to censorship, and even an obligation to censorship, like in NetzDG, GDPR and the Copyright Directive. It's too optimistic to assume that censorship by tech firms would always reckon with old heuristics like "facts are sacred, comments are free" or with the detached ways in which the press quotes or paraphrases subjective sources (Baden & Tenenboim-Weinblatt, 2018; Van Atteveldt, Sheafer, Shenhav, & Fogel-Dror, 2017). A re-allocation of property rights over algorithms should allow stakeholders to bring their cases to court, and should allow judges to make consistent judgements based on vested property rights. "The legislation must refer to the algorithm" (Lodder & Loui, in press). It should be noted that legislation to oblige property owners to give others full access to their records is by no means a new phenomenon. For example, farmers who own land and livestock are obliged by law to give full access to their milk production, manure disposal and use of pesticides. To remove the sting from the Copyright Directive "upload filter" (13), it would be helpful to oblige tech firms to equip firms and citizens like authors and artists with the ownership over advanced monitoring tools to inspect whether their copyright is violated. New fast-track procedures would enable firms, authors and artists to go to court, and would enable judges to shape judicial precedents by commanding tech firms to remove the links to materials with specific features in terms of copyright. This digital copyright should not apply to everything, for example not to news content to which the "link tax" applies, and not to materials whose "linkability", thus whose online availability, occurred with the privity of the artist. To remove the sting from the "link tax" (11), it would be help-

ful to oblige tech firms to equip media companies with the ownership over advanced monitoring tools to register, first, the amount of search queries by internet users that resulted in prominent links to news content delivered by media companies, next to verify that tech firms did not promote news content from media companies who undermine the link tax by not collecting it, and to verify that the link tax is paid directly to the editorial staff of a medium website rather than to the media companies of media tycoons. The link tax should apply only to tech firms with a business model to raise advertisements by attracting users who search for news content, and not, for example to Wikipedia, researchers, teachers or interested citizens. Splitting up the property rights over algorithms of the "semi-benevolent information monopolies" (Karpf, 2018) may reduce automated surveillance and control, and increase the quality of news, ads, and chats.

Acknowledgments

The author is grateful to John Faasse, Wouter van Atteveldt, and Kasper Welbers for comments and suggestions.

Conflict of Interests

The author declares no conflict of interests.

References

Baden, C., & Tenenboim-Weinblatt, K. (2018). Viewpoint, testimony, action. *Journalism Studies*, *19*(1), 143–161. https://doi.org/10.1080/1461670X.2016.1161495

C-Span. (2018). Facebook CEO Mark Zuckerberg hearing on data privacy and protection (Video broadcast). *C-Span*. Retrieved from www.c-span.org/video/?443543-1/facebook-ceo-mark-zuckerberg-testifies-data-protection&utm_source=fark&utm_medium=website&utm_content=link&ICID=ref_fark

Cerf, V., Berners-Lee, T., Esterhuysen, A., Kahle, B., Behlendorf, B., Schneier, B., . . . Cook, W. (2018). *Letter to the President of the European Parliament*. Retrieved from https://www.eff.org/files/2018/06/13/article13letter.pdf

Coase, R. H. (2013). The Federal Communications Commission. *Journal of Law and Economics*, *56*(4), 879–915. (Original work published 1959)

Coase, R. H. (2013). The problem of social cost. *Journal of Law and Economics*, *56*(4), 837–877. (Original work published 1960)

Cook, T. (1998). *Governing with the news*. Chicago, IL: University of Chicago.

Ehle, K. (2018, July 17). European Parliament voted 318 To 278 against the amendment proposal for a copyright directive for the digital single market. *Mondaq*. Retrieved from http://www.mondaq.com/germany/x/719704/Copyright/European+Parliament+Voted+318+to+278+Against+the+Amendment+Proposal+for+a+Copyright+Directive+for+the+Digital+Single+Market

Facebook. (2018). *Hearing: Facebook, social media privacy, and the use and abuse of data* (Memorandum). Washington, DC: Committee on Commerce, Science, and Transportation. Retrieved from www.commerce.senate.gov/public/_cache/files/9d8e069d-2670-4530-bcdc-d3a63a8831c4/7C8DE61421D13E86FC6855CC2EA7AEA7.senate-commerce-committee-combined-qfrs-06.11.2018.pdf

Hobbes, T. (2002). *Leviathan*. Orchard Park, NY: Broadview Press. (Original work published 1660)

Karpf, J. (2018). *Learning to live among semi-benevolent information monopolies* (Working Paper). Prague: International Communication Association.

Kleinnijenhuis, J., Schultz, F., Oegema, D., & van Atteveldt, W. (2013). Financial news and market panics in the age of high-frequency sentiment trading algorithms. *Journalism*, *14*(2), 271–291. https://doi.org/10.1177%2F1464884912468375

Kräussl, R., & Mirgorodskaya, E. (2017). Media, sentiment and market performance in the long run. *The European Journal of Finance*, *23*(11), 1059–1082. https://doi.org/10.1080/1351847X.2016.1226188

Lee, J., & Xu, W. (2018). The more attacks, the more retweets: Trump's and Clinton's agenda setting on Twitter. *Public Relations Review*, *44*(2), 201–213. https://doi.org/10.1016/j.pubrev.2017.10.002

Lodder, A. R., & Loui, R. (in press). Data algorithms and privacy in surveillance: On stages, numbers and the human factor. In W. Barfield & U. Pagallo (Eds.), *Research handbook of law and artificial intelligence*. Cheltenham: Edward Elgar.

Newman, N., Fletcher, R., Kalogeropoulos, A., Levy, D. A., & Nielsen, R. K. (2018). *Reuters Institute digital news report 2017*. Oxford: Reuters Institute.

Thurman, N., & Schifferes, S. (2012). The future of personalization at news websites. *Journalism Studies*, *13*(5/6), 775–790. https://doi.org/10.1080/1461670X.2012.664341

UK Music. (2018, July 3). UK Music chief slams Google as "corporate vultures" as figures show Google's €31m EU lobbying bid. *UK Music*. Retrieved from https://www.ukmusic.org/news/uk-music-chief-slams-google-as-corporate-vultures-as-figures-show-googles-3

Van Atteveldt, W., Sheafer, T., Shenhav, S. R., & Fogel-Dror, Y. (2017). Clause analysis: Using syntactic information to automatically extract source, subject, and predicate from texts with an application to the 2008–2009 Gaza War. *Political Analysis*, *25*(2), 207–222. https://doi.org/10.1017/pan.2016.12

Wells, C., Shah, D. V., Pevehouse, J. C., Yang, J., Pelled, A., Boehm, F., . . . Schmidt, J. L. (2016). How Trump drove coverage to the nomination: Hybrid media campaigning. *Political Communication*, *33*(4), 669–676. https://doi.org/10.1080/10584609.2016.1224416

Zuckerberg, M. [Mark]. (2018, March 21). I want to

share an update on the Cambridge Analytica situation—including the steps we've already taken and our next steps to address this important issue [Facebook status update]. Retrieved from https://www.facebook.com/zuck/posts/10104712037900071

About the Author

Jan Kleinnijenhuis (PhD Vrije Universiteit Amsterdam, 1990) is a Professor of Communication Science at the Vrije Universiteit Amsterdam. His research addresses the nature of political and economic news, news selection and agenda building, news effects, and methods for automated and semi-automated content analysis, or, text mining.

COGITATIO

Media and Communication (ISSN: 2183–2439)
2018, Volume 6, Issue 3, Pages 83–92
DOI: 10.17645/mac.v6i3.1499

Article

Psychopaths Online: The Linguistic Traces of Psychopathy in Email, Text Messaging and Facebook

Jeffrey T. Hancock [1,*], Michael Woodworth [2] and Rachel Boochever [3]

[1] Department of Communication, Stanford University, Stanford, 94305, USA; E-Mail: jeff.hancock@stanford.edu
[2] Department of Psychology, University of British Columbia, Kelowna, V1V 1V7, Canada; E-Mail: michael.woodworth@ubc.ca
[3] Law School, Stanford University, Stanford, 94305, USA; E-Mail: rcb@stanford.edu

* Corresponding author

Submitted: 27 March 2018 | Accepted: 19 September 2018 | Published: 28 September 2018

Abstract
Individuals high in psychopathy are interpersonally manipulative, exhibit callous affect, and have criminal tendencies. The present study examines whether these attributes of psychopathy are correlated with linguistic patterns present in everyday online communication. Participants' emails, SMS messages, and Facebook messages were collected and analyzed in relation to their scores on the Self-Report Psychopathy Test III. The findings suggest that psychopathic tendencies leave a trace in online discourse, and that different forms of online media sometimes moderate the association between a linguistic dimension and psychopathy scores. Consistent with previous studies and the emotional and interpersonal deficits central to psychopathy, participants higher in psychopathy showed more evidence of psychological distancing, wrote less comprehensible discourse, and produced more interpersonally hostile language. The results reveal that linguistic traces of psychopathy can be detected in online communication, and that those with higher traits of psychopathy fail to modify their language use across media types.

Keywords
computer-mediated communication; email; Facebook; online media; psychopathy; text messaging

Issue
This article is part of a multidisciplinary issue of *Media and Communication*, edited by Epp Lauk (University of Jyväskylä, Finland) and Raul Reis (Emerson College, USA).

1. Introduction

Language produced during communication can be indicative of personality traits, as well as emotional states, identity and cognitive style (Pennebaker, 2011; Pennebaker & King, 1999; Pennebaker, Mehl, & Niederhoffer, 2003). As Pennebaker (2011) describes, the words we use are like fingerprints; we all leave traces of ourselves behind in our words. A series of studies has examined the relationship between language use and the personality construct of psychopathy; more specifically, whether discourse patterns reflect psychopathic tendencies (Hancock, Woodworth, & Porter, 2013). For example, people high in psychopathy show evidence of narcissism and psychological distancing in their language (Hancock et al., 2012).

While these results are important, they are limited in their generalizability due to narrow populations (e.g., imprisoned murderers) and limited discourse types (e.g., narratives elicited by researchers). Pennebaker (2011) notes that language styles are adaptive based on different situations. Thus, natural language, or language produced in the real world, would arguably provide better insight into one's personality. The increasing popularity of online communication, with its automatic storage capabilities, means that it is now easy to access language that has been produced naturally, enabling analysis of real world language. A good deal of research on

online communication also suggests that these media affect how people communicate (e.g., Herring & Androutsopoulos, 2015). This exploratory study seeks to examine the relationship between psychopathy and natural language produced in online communication. We also examine whether language produced in online communication is fundamentally different from language elicited in a laboratory setting.

1.1. Attributes of Psychopathy

The concept of psychopathy was first officially documented by Hervey Cleckley (1941/1976), who documented a subsection of his patients that displayed a dire lack of morality. The current conceptualization of a psychopath is of a manipulative, cunning and antisocial individual who, according to Hare (2006), comprise about 1% of the general population. Although people with psychopathic tendencies lack key emotional traits such as remorse and guilt, and generally present shallow affect, they have an uncanny ability to influence and deceive others. As mentioned above, previous studies have suggested that psychopaths' language is distinct from that of non-psychopaths, but additional studies are needed to examine the exact linguistic mechanisms that may be contributing to their success in interactions with others.

1.2. Psychopathy and Language

Several previous studies have used automatic linguistic analysis to examine the discourse of psychopaths convicted of murder. In one study (Hancock et al., 2013), psychopathic murderers, as opposed to non-psychopathic murderers, used more subordinating conjunctions (*because*, *since*, *as*, *so that*, etc.), evidencing of the use of more cause and effect statements, and indicating that the psychopaths' viewed their murders as more instrumental than non-psychopaths (see Woodworth & Porter, 2002). Psychopathic murderers used more past tense words, fewer present tense words, more articles, and more concrete nouns than non-psychopathic murderers. The combination of these discourse patterns suggests that the psychopaths were more psychologically distanced from these incidents, likely reflecting their emotional deficits even when recounting a negative event as extreme as homicide. Psychopathic murderers also used twice as many content words relating to basic physiological needs, such as *food*, *money*, and *shelter*, but fewer words relating to higher interpersonal needs, such as *family* and *love*. This combination suggests that psychopaths are less concerned with meaningful relationships and more focused on basic needs and achievement of goals (see Woodworth & Porter, 2002, and Porter & Woodworth, 2007, for research considering the instrumental nature of psychopaths specifically for the offence of homicide). Finally, psychopaths produced more disfluencies, such as *uh* and *um*, when describing their murder than non-psychopaths, suggesting a less fluent and comprehensible narrative.

Other work has focused more on the language comprehension skills of psychopaths, which may have some bearing on our examination of their language production. For instance, Vaughn et al. (2011) found that juveniles high in psychopathy had poor reading achievement. Poor reading achievement may be reflective of the emotional deficits observed in psychopaths. Do people with high psychopathic tendencies also produce less readable language?

1.3. Online Communication and Psychopathy

The present study aims to expand the research discussed above by examining whether language produced during everyday online communication contains patterns associated with psychopathy. One of the most important attributes of online communication is that it leaves digital traces of messages, which allows researchers to examine real messages exchanged during the course of everyday communication. Compared to narratives elicited for an experiment, naturally produced language in online communication should be less formal, less prone to social desirability motivations, and reflect more realistic interpersonal dynamics. For instance, a number of recent studies suggest that private traits are predictable from records taken from online communication (e.g., Bachrach, Kosinski, Graepel, Kohli, & Stillwell, 2012; Kosinski, Stillwell, & Graepel, 2013). Given these attributes of online communication, we first expect that language in online communication will be different than experimentally elicited narratives, and second that more correlations between language dimensions and psychopathy will be observed in online communication than in elicited narratives.

With respect to the specific linguistic dimensions and attributes of psychopathy, we expect to find similar evidence of narcissistic tendencies and psychological distancing in online communication as found in previous studies (Hancock et al., 2012; Hancock et al., 2013; Sumner, Byers, Boochever, & Park, 2012). Narcissism has been examined extensively in the context of social media (e.g., Carpenter, 2012; Davenport, Bergman, Bergman, & Fearington, 2014), and based on these studies we predict that participants higher in psychopathy will use more first person singular pronouns, but fewer first person plural pronouns and fewer second person pronouns, and they should display increased psychological distancing in their language.

Reflecting the interpersonally manipulative characteristic of psychopathy, we expect that participants higher in psychopathy will use more cause and effect statements (Woodworth & Porter, 2002). In particular, participants higher in psychopathy should use more subordinating conjunctions (*because*, *so*, etc.). We also expect to see evidence of psychopaths' increased focus on basic needs (*food*, *drink*, *money*, etc.) and reduced focus on higher level needs (*spirituality*, *relationships*, *religion*, etc.), consistent with Hancock et al. (2013).

Given prior work showing that juveniles high in psychopathy have poor reading achievement (Vaughn et al., 2011), and that psychopaths tend to be more disfluent during speech (Hancock et al., 2013), we predict that participants higher in psychopathy will produce less comprehensible language as measured by a standard readability index.

We expect participants higher in psychopathy to produce more words associated with a hostile interpersonal demeanor, such as anger terms and swear words. Although this may seem contradictory considering their profound emotional deficits, another characteristic of psychopathy is poor behavioral control, which refers specifically to their quick temper in some situations. This poor behavioral control should be reflected in their language to a larger degree than participants scoring lower in psychopathy, specifically for conversations discussing a situation that was contentious. Indeed, recent work examining the dark triad in Twitter discourse (Sumner et al., 2012) found that users high in psychopathy and Machiavellianism used more swear words and more words associated with anger than other users. This finding was recently replicated by Bogolyubova, Panicheva, Tikhonov and Ivanov (2018), who found that male Facebook users scoring high in psychopathy sent more aggressive and insulting posts than users low in psychopathy. Thus, participants higher in psychopathy should use more words associated with anger and swear words in online communication.

Psychopathy can be assessed along several dimensions. For example, the Self-Report Psychopathy Scale III (SRP-III; Paulhus, Hemphill, & Hare, 2012) conceptualizes the psychopathy with a four-factor structure, consistent with the Psychopathy Checklist-Revised (Mahmut, Menictas, Stevenson, & Homewood, 2011). This measure of psychopathy includes: 1) Interpersonal Manipulation, the degree to which an individual finds it easy to manipulate or deceive others, 2) Criminal Tendencies, the degree to which an individual has been involved in illegal activities, 3) Erratic Lifestyle, the degree to which an individual is prepared to violate social norms and agreements, and 4) Callous Affect, the degree to which an individual displays a lack of empathy for others. An important research question is whether these four factors may be related to different language patterns in online communication. For example, hostile and aggressive language may be most associated with interpersonal manipulation, while psychological distancing might be more related to callous affect. How does language in online communication relate to these sub-factors of psychopathy?

Online communication is also not monolithic. The three types of online communication examined here may differentially reveal associations between language and psychopathy. As is now well-known in the communication and technology literature, different media types have unique sets of affordances that shape communication patterns (Clark & Brennan, 1991; Whittaker, 2002). For example, SMS text messaging tends to be less formal than email, while Facebook messages tend to be more public than either SMS or email (Bazarova, 2012). How will these characteristics of each media affect the relationship between psychopathy and language?

Finally, our last research question was concerned with whether these language features could be used to classify people as scoring low or high on psychopathy. Prior work has found limited success in classifying Twitter users with machine learning techniques. For example, Sumner et al. (2012) found that the best models performed above chance, but that the models overall had poor accuracy when classifying individuals as scoring high in psychopathy. Here we examine whether having records from a variety of online communication and from elicited narratives can improve classification performance compared to the short messages found in Twitter discourse used in prior studies.

2. Methods

2.1. Participants

The participants in this study were 110 undergraduate students at a large US research university. The participants were recruited using the university's psychology experiment recruitment tool. They received course credit or $5 in compensation for participating in this study. Of the 110 participants, 85 were female. Ages ranged from 18 to 24 years old, with a mean age of 20.2 years ($std = 1.28$).

2.2. Materials

We used the SRP-III (Paulhus et al., 2012) to measure psychopathic tendencies. The SRP-III is a reliable self-report measure of psychopathy that has demonstrated both convergent and divergent validity with a four-factor structure consistent with the Psychopathy Checklist-Revised (Mahmut et al., 2011). The SRP-III has been tested using both student samples and community samples, finding similar results (Williams, Nathanson, & Paulhus, 2003; Williams, Paulhus, & Hare, 2007; Mahmut et al., 2011). It consists of 64 questions, 16 questions each relating to four facets of psychopathy: callous affect (CA; e.g., "I am often rude to other people"), erratic lifestyle (ELS; e.g., "Rules are made to be broken"), interpersonal manipulation (IM; e.g., "I find it easy to manipulate people"), and criminal tendency (CT; e.g., "I have broken into a building or vehicle to steal or vandalize"). The higher the SRP-III score across these four factors the higher the psychopathic tendency for that participant.

The SRP-III was scored based on instructions from Paulhus et al.'s *Manual for the Self-Report Psychopathy Scale*. Each of the 64 questions was answered on a Likert scale from 1 (strongly disagree) to 5 (strongly agree). Of the 64 questions, 21 items were reverse-coded. All answers were summed to create a total SRP-III score, with a total possible range of 64 to 320, and a possible range of

16 to 80 for each of the four facets. For the present sample, the SRP-III mean total was 143.1 (SD = 27.4), with a minimum score of 82 and a maximum of 215.

For the elicited narratives, participants were asked to write two short stories, one about a positive event in their life, and one about a negative event in their life. They were asked to write stories of approximately 100 words. The negative elicited stories on average were 116 words (SD = 36.3) while the positive stories were 110.0 words (SD = 33.9). For the online communication messages, the average for text messages was 140.8 (SD = 78.5) words, for Facebook messages 142.5 (SD = 96.1) words, and for email was 700.2 (SD = 397.1) words.

2.3. Procedure

This study was conducted in the form of an online web survey via Qualtrix. After providing informed consent, participants were asked to complete the SRP-III and submit four types of language samples: 1) a short positive story and a short negative story, each about 100 words, which we refer to as "elicited narratives", 2) their twenty most recently sent SMS messages, 3) their twenty most recent sent emails, and 4) five each of their most recent Facebook status updates, private messages, and wall posts. For the elicited narratives, participants typed the stories into a text box. For the SMS messages, emails, and Facebook messages, collectively called "online communication" samples, participants transcribed their SMS messages from their mobile phone into the survey, and for the email and Facebook samples participants simply copied them from a web browser into the survey form. Participants were also asked demographic questions at the end of the survey (age and gender). Once they completed the survey they received debriefing information and remuneration.

2.4. Text Analysis

All language samples were converted into word documents and analyzed using the text analysis program *Linguistic Inquiry and Word Count* (LIWC; Pennebaker, Boothe, & Francis, 2007). LIWC identifies and categorizes words based on linguistic dimensions, psychological constructs, personal concern categories, and paralinguistic dimensions, among other output variables (Hancock et al., 2012). Based on a dictionary of almost 4,500 words and word stems, LIWC counts the amount of words in each category, and divides the sum by the word count. This gives the percentage each category represents of the total word count, normalizing for verbosity. LIWC has been validated in a large number and variety of psychology and communication studies (Pennebaker, 2011).

To analyze the attributes of psychological distancing, basic needs, high-level needs, and readability we used algorithms of combined multiple standardized linguistic variables from LIWC. For psychological distancing, we used Chung & Pennebaker's (2007) algorithm, a combination of: six letter words, articles, past tense, and the inverse of first person singular pronouns, present tense and discrepancy words. For basic and higher level needs, we used sets of words identified in previous studies and mapped on to the LIWC dictionary (Hancock et al., 2012; Woodworth & Porter, 2002): basic needs include sex, money, leisure, achievement, work, health, and biology in the LIWC dictionary; higher level needs include family, religion, positive emotion, social, friends, and the inverse of death. Lastly, we based readability on an approximation of the Flesch-Kincaid readability test (Kincaid, Fishburn, Rogers Jr., & Chissom, 1975), calculated by multiplying negative one by six letter words, adding words per sentence, subtracting the amount of words recognized by the LIWC dictionary, and adding three. Finally, we included word categories in LIWC that reflect aggressive discourse, including swear words (e.g., "asshole") and anger terms (e.g., "dammit").

3. Results

3.1. Data Analysis

Our analytic approach involved two steps. First we used linear mixed models to test for language differences between discourse type (online communication vs. elicited narratives), with the two discourse types nested within each participant to account for non-independence between the discourse samples. This analysis revealed significant differences between online communication and elicited narratives for virtually all of the language dimensions. Thus, in the second step of the analysis we examined online communication and elicited narratives separately. In this second step we calculated the bivariate correlations between each language variable and the SRP-III factors and used linear mixed models to examine possible interactions with media type (email, SMS text messaging and Facebook) for online communication and valence (positive vs. negative) for elicited narratives.

3.2. Discourse Patterns in Online Communication versus Elicited Narratives

Our first research question examined whether the language patterns in online communication were different from those observed in elicited narratives along the primary dimensions of interest. We combined the email, Facebook, and SMS transcripts to create an online Communication Index for each language variable, and we combined the positive and negative narratives to create an Elicited Narrative Index.

As can be seen in Table 1, significant differences between online communication and elicited narrative discourse emerged for most language variables, including pronouns, verb tense, and emotions. In elicited narratives, participants wrote more about themselves, wrote more about past events, and used fewer emotion words

Table 1. Mean (*SD*) discourse features across online communication and elicited narratives as a percentage of total words.

		Online Communication		Elicited Narratives			
		M	*SE*	*M*	*SE*	*F*(1,436)	P
Pronouns	I	6.61%	.19%	9.47%	.22%	126.9	<.001
	We	1.03	.09	1.21	.10	1.95	.14
	You	3.86	.11	.05	.13	623.4	<.001
Tense	Past	2.64	.14	8.95	.17	999.5	<.001
	Present	12.41	.21	3.31	.24	1012.7	<.001
Emotion	Positive	7.24	.21	3.88	.25	112.3	<.001
	Negative	1.67	.11	2.25	.13	11.4	.001
	Anger	.47	.05	.37	.06	1.7	.19
	Swear	.25	.03	.02	.04	24.1	<.001

than in their online discourse. In general, these results suggest that the elicited narratives were focused more on the participants' actions in the past, and despite being asked to write about positive and negative experiences, these stories were overall less emotional than their everyday online discourse. It is also important to note that participants used more words relating to anger and more swear words in online communication than in elicited narratives, indicating that language use in online communication is more colloquial and informal than what participants write in a laboratory setting. These data support our expectations that discourse patterns in online communication differ substantially from elicited narratives.

3.3. Psychopathy Scores and Online Communication

Our primary question of interest was how online communication predicts psychopathy scores. Table 2 describes the simple bivariate correlations for the language dimensions from the combined Online Communication Index and psychopathy scores including the SRP-III sub-factors. To compare across the three media types (SMS text messaging, email, Facebook) we used linear mixed models to account for the non-independence of the writing samples for each participant. A mixed model was created for each language variable, with media type (email, SMS, Facebook) nested within participant, to predict the overall SRP-III psychopathy score. For each hypothesis, the mixed model tested the effect of the language variable on psychopathy scores and the interaction between the media type and the language variable (whether the association between psychopathy scores and the language variable was different across media type).

3.3.1. Narcissistic Tendencies

We expected that participants higher in psychopathy would exhibit narcissistic tendencies in the pattern of their pronoun use, with increased focus on self and decreased focus on others. This hypothesis was partially supported. As can be seen in Table 2, second person pronoun pronouns were negatively correlated with overall psychopathy scores, suggesting that more psychopathic individuals referred less often to other people in their conversations. This effect was driven by negative associ-

Table 2. Pearson bivariate 2-tailed correlations between SRP-III scores and selected linguistic factors in online communication.

	SRP-III Total	Interpersonal Manipulation	Callous Affect	Erratic Life Style	Criminal Tendency
I	.05	.01	.03	.09	.04
We	−.02	.01	−.02	−.03	−.01
You	−.24*	−.17	−.21*	−.28**	−.12
Psychological Distancing	.19*	.11	.24*	.23*	.05
Conjunctions	−.12	−.14	−.19*	.10	−.18
Basic Needs	−.02	−.15	−.03	.12	−.01
Higher Needs	−.12	−.14	−.09	−.08	−.08
Readability	−.32**	−.24*	−.22*	−.34**	−.23*
Anger	.22*	.23*	.13	.18	.17
Swear	.31**	.30**	.23*	.21*	.25**

Notes: * Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed).

ations with the Callous Affect and Erratic Life Style factors, indicating that participants that scored higher on these two factors were less likely to refer to their communication partner (e.g., *you*). The mixed model analysis revealed that the inverse relationship between second person pronoun use and psychopathy scores was not affected by the media type, $F(1,322) < 1$. Thus, increased levels of psychopathy were associated with reduced second person pronoun use across all three forms of online communication.

First person singular (*I*) and plural (*we*), however, which were expected to correlate positively with psychopathy scores, were not associated with the SRP-III scores (see Table 2). Thus, while participants with higher psychopathy scores referred to other people less in their online conversations, suggesting reduced attention towards their conversation partner, they did not also explicitly focus more on themselves.

3.3.2. Psychological Distancing

We expected that participants higher in psychopathy would use more psychological distancing. This hypothesis was supported in the correlation analysis. Psychological distancing was negatively associated with overall psychopathy scores. This effect was driven by negative correlations with Callous Affect and Erratic Lifestyle, suggesting that participants scoring higher on these factors wrote with more psychological distance in their online communication.

The linear mixed model analysis produced a significant interaction, however, between media type and psychological distancing, $F(1,321) = 3.42$, $p < .034$. This interaction revealed that the association between psychological distancing was most evident in email ($b = .27$, $p < .005$), followed by language from Facebook ($b = .19$, $p < .05$), but was not significant in Facebook posts ($b = -.05$, *ns*). These data suggest that the type of online communication plays an important role in how the psychological distancing associated with psychopathy is expressed in language use.

3.3.3. Subordinating Conjunctions

We expected that participants scoring high in psychopathy would use more conjunctions given prior work indicating that psychopaths tend to view the world instrumentally. This hypothesis was not supported. In fact, in contrast with our expectations, participants high in the callous affect factor of psychopathy used fewer subordinating conjunctions. We did not explore this unexpected correlation further.

3.3.4. Basic vs. High-Level Needs

Our hypothesis that participants high in psychopathy would focus more on basic needs and less on higher level needs was not supported. No effects were observed between psychopathy scores and either basic or higher level needs in online discourse.

3.3.5. Readability

Readability was negatively correlated with total SRP-III scores and all four sub-factors in online discourse. The mixed model analysis, however, produced a marginally significant interaction between media type and readability in predicting psychopathy, $F(1,320) = 2.67$, $p = .07$. The interaction revealed that the association between readability and psychopathy was only evident in email discourse ($b = -.28$, $p < .003$), and not in Facebook or SMS text messaging. One possible reason is that the range of readability of Facebook and SMS text messages is constrained by word length limitations (e.g., SMS text messaging is typically 140 characters).

3.3.6. Hostile Interpersonal Style

Our hypothesis that participants scoring higher in psychopathy would use more words associated with anger and swear words was supported. The frequency of swear words correlated with the total psychopathy score and each sub-factor, and anger words correlated with the total score and interpersonal manipulation. The linear mixed model analysis revealed no interaction effect with media type for either anger words or swear words, suggesting that this hostile interpersonal style is evident across the three media types for participants scoring highly on the SRP-III.

3.4. Psychopathy Scores and Elicited Narratives

The same analysis was conducted on the elicited narratives. The bivariate correlations are reported in Table 3. To compare across the valence of the elicited narratives (positive vs. negative) we again used linear mixed models to account for the non-independence of the writing samples for each participant. A mixed model was created for each language variable, with narrative valence nested within participant, to predict the overall SRP-III psychopathy score.

Unlike the frequent associations between psychopathy scores and online communication, the language from the elicited narratives were correlated with psychopathy scores along only two dimensions. Consistent with our hypothesis, participants scoring higher in psychopathy wrote their narratives with more words related to basic needs, and this association was primarily driven by scores on criminal tendency. Also as expected, participants scoring higher in psychopathy wrote narratives that were lower in readability, which was driven by the callous affect and criminal tendency factors. The mixed model analysis revealed that there were no interaction effects between media type and language dimension, suggesting the basic needs and readability effects were the same for positive and negative narratives.

Table 3. Pearson bivariate 2-tailed correlations between SRP-III scores and selected linguistic factors in elicited narratives.

	SRP-III Total	Interpersonal Manipulation	Callous Affect	Erratic Life Style	Criminal Tendency
I	.01	.05	.04	−.03	−.04
We	.03	.05	−.04	.06	.04
You	−.09	−.01	−.08	−.11	−.11
Psychological Distancing	.08	.06	.09	.06	. 07
Conjunctions	−.07	−.03	−.07	−.05	−.07
Basic Needs	.14*	−.04	.07	.13	.22**
High-Level Needs	−.04	−.03	−.02	.01	−.07
Readability	−.15*	−.08	−.15*	−.09	−.17*
Anger	−.03	−.02	−.02	−.03	−.04
Swear	−.04	.01	−.01	−.13	.02

Notes: * Correlation is significant at the 0.05 level (2-tailed); ** Correlation is significant at the 0.01 level (2-tailed).

3.5. Classifying Low versus High on the SRP-III.

Our last question was concerned with how the language features in online communication could be used to classify participants as scoring high versus low in psychopathy. To address this question we conducted a quartile split on the SRP-III data, with participants scoring below 123 on the scale in the bottom quartile (Low Psychopathy) and those scoring above 162 in the top quartile (High Psychopathy). A logistic regression predicting Low vs. High Psychopathy was constructed by entering each of the language features that displayed bivariate correlations with scores on the SRP-III (see Table 2), including second person pronouns, anger terms, swear words, readability and psychological distancing. The regression was significant, χ^2 (5) = 42.32, $p < .001$, and accounted for approximately 36.7% of the variance. The classification accuracy was 78.9%, with 16 of the 22 high scorers correctly identified and 17 of the 23 low scorers correctly identified.

4. Discussion

The goal of this study was two-fold: 1) to examine discourse patterns associated with psychopathy in two different communication contexts, online communication and elicited narratives, and 2) to examine how psychopathy is expressed differently across several types of online communication. Language collected from archived emails, SMS text messages, and Facebook messages revealed that language produced in online communication was significantly different than language elicited for the purpose of a study in terms of pronoun use, verb tense, and emotion terms. In addition, more correlations between various components of psychopathy were found with language produced naturally in online communication than in the elicited narratives, suggesting online discourse is a rich source of communication that can reveal key aspects of the self.

Several hypotheses about the association between the SRP-III psychopathy scores and linguistic patterns in online communication were supported. In online communication participants higher in psychopathy referred less often to their conversation partner, used more psychological distancing, produced less comprehensible text, and used more interpersonally hostile language, such as anger and swear words. However, participants higher in psychopathy did not focus more on basic needs or less on higher level needs in online communication. In contrast, a positive correlation with basic needs and psychopathy scores was observed in the elicited narratives, one of the few instances where associations emerged in the narratives but not in online communication. For instance, narratives are more likely to provide insight for why a person performed an action (e.g., "At the birthday party I just wanted to eat the cake because I was starving"). If this is the case then research examining the role that motivations, such as basic needs, play in psychopathy should focus on elicited narratives rather than naturally produced communication. Finally, psychopathy scores were not correlated with conjunctions, suggesting that participants higher in psychopathy did not use more cause and effect statements.

Despite important differences between the criminal and student samples, evidence of psychological distancing in the current study demonstrates that some aspects of Hancock et al.'s (2013) findings with criminal psychopaths may also apply to a more general population. Psychological distancing suggests that psychopaths do not emotionally connect with what they are saying and that they are either detached from their language or use the same type of language to refer to both emotional and non-emotional concepts. This was particularly the case for individuals scoring high on Callous Affect and Erratic Lifestyle, which both suggest deficiencies in social and interpersonal functioning.

Additionally, it is important to understand language production in relation to the reading comprehension

deficits observed with people with psychopathic tendencies (Vaughn et al., 2011). The finding that participants higher in psychopathy produced less comprehensible text in both online communication and elicited narratives is consistent with previous studies, such as Brinkley, Newman, Harpur and Johnson (1999), and demonstrates for the first time that some of the speech difficulties found in previous studies, such as lack of cohesion or increased speech disfluencies, are also apparent in text-based online communication contexts in the form of reduced readability.

Evidence of increased anger and swearing in online communication is consistent with psychopath's proclivity for interpersonal manipulation, as well as the poor behavioral controls associated with callous affect, specifically for anger. For example, one participant who scored a 200 on the SRP-III (higher than two standard deviations above the mean score) wrote in an email: "I do not wish to talk to you anymore about anything ever again. I'm glad that this is over because talking to you is like sticking a spoon in my ass". The same participant had the following Facebook status updates: "Dead", "Bored", "Tired", "Fighting with her again", and "Hate everyone". While this person may be an outlier given their very high SRP-III score, these aggressive and insulting posts are consistent with prior research finding that people scoring high on psychopathy tend to send messages with more negative emotion terms and swear words (Bogolyubova et al., 2018; Sumner et al., 2012).

It is also important to note that anger and swearing are significantly correlated with SRP-III scores in online communication, but not in the elicited narratives. Psychopaths are known for their impulsivity and their increased usage of swear words and anger words could indicate their reduced ability to control the *type* of language they are producing (e.g., negative) in natural discourse. The finding that they are even less able to do this in online communication contexts warrants further investigation. For instance, we found that the swear words and anger terms were correlated with psychopathy scores in one-to-one forms of communication, including email and text messaging, but not in broadcast forms of communication, such as posting on Facebook. These data suggest that participants higher in psychopathy managed to keep their hostile language limited to direct communication and avoided using it in more public online communication. More work is required to understand how the features of online communication specifically trigger or exacerbate a psychopath's impulsive nature, as may be the case for online trolls (Lopes & Yu, 2017).

Despite the intriguing findings outlined above, this study is limited by both the small sample size, $N = 110$, and the small size of the language samples collected. A larger, more diverse group of participants and a greater number of language samples could produce different results. Nonetheless, these data provided sufficient power to observe predicted results along many of the hypotheses. An important direction to build on this work, however, is attempting to include more individuals that score highly on the SRP-III. Psychopathy scores were primarily in the low and moderate ranges, consistent with the rates of psychopathy in general community samples, with few reaching the actual cut-off denoting psychopathy. The fact that such differences were found across the range of psychopathy scores emphasizes the importance of our findings considering the truncated range of psychopathy in the current sample.

5. Conclusion

This study extended prior work examining the discourse patterns of psychopaths and non-psychopaths by examining different forms of online discourse. Our findings, across three types of online communication (email, Facebook and SMS text messaging) support previous research, showing that discourse patterns of participants' higher in psychopathy showed evidence of narcissism, psychological distancing, produced less comprehensible text, and used more words indicative of an interpersonally hostile style, including more anger and swear words. These results were more pronounced in online discourse than in elicited narrative discourse, suggesting that real world discourse is more revealing of psychopathic tendencies. There may be features unique to online communication that afford a better opportunity to spot these linguistic traces of psychopathy, or online interactions that are more likely to trigger or prompt these differences. Regardless, our results reinforce the theory that individual personality characteristics, such as psychopathic tendencies, can be reflected in discourse patterns found in online communication.

Acknowledgments

This research was funded in part by a grant from SSHRC awarded to the second author. The authors also acknowledge Sunny Liu for comments on a previous draft.

Conflict of Interests

The authors declare no conflict of interests.

References

Bachrach, Y., Kosinski, M., Graepel, T., Kohli, P., & Stillwell, D. (2012). Personality and patterns of Facebook usage. In *Proceedings of Web Science '12* (pp. 24–32). Evanston, IL: ACM.

Bazarova, N. N. (2012). Public intimacy: Disclosure interpretation and social judgments on Facebook. *Journal of Communication*, *62*(5), 815–832.

Bogolyubova, O., Panicheva, P., Tikhonov, R., Ivanov, V., & Ledovaya, Y. (2018). Dark personalities on Facebook: Harmful online behaviors and language. *Computers in Human Behavior*, *78*, 151–159.

Brinkley, C. A., Newman, J. P., Harpur, T. J., & Johnson, M. M. (1999). Cohesion in texts produced by psychopathic and non-psychopathic criminal inmates. *Personality and Individual Differences*, *26*(5), 873–885.

Carpenter, C. J. (2012). Narcissism on Facebook: Self-promotional and anti-social behavior. *Personality and Individual Differences*, *52*(4), 482–486.

Chung, C. K., & Pennebaker, J. W. (2007). The psychological function of function words. In K. Fiedler (Ed.), *Social communication: Frontiers of social psychology* (pp. 343–359). New York, NY: Psychology Press.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), *Perspectives on socially shared cognition* (pp. 127–149). Washington, DC: American Psychological Association.

Cleckley, H. (1976). *The mask of sanity* (5th ed.). St. Louis, MO: Mosby. (Original work published 1941)

Davenport, S., Bergman, S. M., Bergman, J. Z., & Fearrington, M. E. (2014). Twitter versus Facebook: Exploring the role of narcissism in the motives and usage of different social media platforms. *Computers in Human Behavior*, *32*, 212–220.

Hancock, J. T., Woodworth, M., & Porter, S. (2013). Hungry like the wolf: A word pattern analysis of the language of psychopaths. *Legal and Criminological Psychology*, *18*(1), 102–114.

Herring, S. C., & Androutsopoulos, J. (2015). Computer-mediated discourse 2.0. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The handbook of discourse analysis* (2nd ed., pp. 127–151). West Sussex: John Wiley & Sons.

Hare, R. D. (2006). Psychopathy: A clinical and forensic overview. *Psychiatric Clinics of North America*, *29*(3), 709–724.

Kincaid, J. P., Fishburn, R. P., Rogers Jr., R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas for Navy enlisted personnel* (Research Branch Report, 8–75). Naval Air Station Memphis, Millington, Tennessee.

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Science*, *110*(15), 5802–5805.

Lopes, B., & Yu, H. (2017). Who do you troll and why: An investigation into the relationship between the dark triad personalities and online trolling behaviours towards popular and less popular Facebook profiles. *Computers in Human Behavior*, *77*, 69–76.

Mahmut, M. K., Menictas, C., Stevenson, R. J., & Homewood, J. (2011). Validating the factor structure of the self-report psychopathy scale in a community sample. *Psychological Assessment*, *23*(3), 670–678.

Paulhus, D. L., Hemphill, J. D., & Hare, R. D. (2012). *Manual for the Self-Report Psychopathy Scale*. Toronto: Multi-Health Systems.

Pennebaker, J. W. (2011). *The secret life of pronouns: What our words say about us*. New York, NY: Bloomsbury Press.

Pennebaker, J. W., Boothe, R. J., & Francis, M. E. (2007). *Linguistic inquiry and word count*. LIWC [Computer software]. Austin, TX: liwc. net.

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, *77*(6), 1296–1312.

Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual Review of Psychology*, *54*(1), 547–577.

Porter, S., & Woodworth, M. (2007). "I'm sorry I did it…but he started it": A comparison of the official and self-reported homicide descriptions of psychopaths and non-psychopaths. *Law and Human Behavior*, *31*(1), 91–107.

Sumner, C., Byers, A., Boochever, R., & Park, G. J. (2012). Predicting dark triad personality traits from twitter usage and a linguistic analysis of tweets. In *Machine learning and applications (icmla), 2012 11th international conference* (Vol. 2, pp. 386–393). IEEE.

Vaughn, M. G., DeLisi, M., Beaver, K. M., Wexler, J., Barth, A., & Fletcher, J. (2011). Juvenile psychopathy personality traits are associated with poor reading achievement. *Psychiatric Quarterly*, *82*(3), 177–190.

Whittaker, S. (2002). Theories and methods in mediated communication. In A. Graesser, M. Gernsbacher, & S. Goldman, (Eds.), *The handbook of discourse processes* (pp. 243–286). Hillsdale, NJ: Erlbaum.

Williams, K. M., Nathanson, C., & Paulhus, D. L. (2003). Structure and validity of the Self-Report Psychopathy Scale-III in normal populations. Presented at the *111th annual convention of the American Psychological Association*. Toronto, Canada.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment*, *88*(2), 205–219.

Woodworth, M., & Porter, S. (2002). In cold blood: Characteristics of criminal homicides as a function of psychopathy. *Journal of Abnormal Psychology*, *111*(3), 436–445.

About the Authors

Jeffrey T. Hancock is a Professor in the Department of Communication at Stanford University. His work is concerned with how social media affect psychological and interpersonal processes, with a particular emphasis on understanding how language can reveal psychological and social dynamics.

Michael Woodworth is a Professor in the Department of Psychology at University of British Columbia, Okanagan. He is a clinical psychologist with research interests in forensic psychology, psychopathy, deception detection, mental health social media and computer-mediated communication research.

Rachel Boochever graduated from Cornell University with a BS in Communication with distinction in research. She graduated from Stanford Law School in 2018 and is an Associate at Wilson Sonsini Goodrich & Rosati in Palo Alto.

Media and Communication (ISSN: 2183-2439)

Media and Communication is an international open access journal dedicated to a wide variety of basic and applied research in communication and its related fields. It aims at providing a research forum on the social and cultural relevance of media and communication processes.

www.cogitatiopress.com/mediaandcommunication

COGITATIO