The algorithmic persuasion framework in online communication: conceptualization and a future research agenda

Zarouali, B.; Boerman, S.C.; Voorveld, H.A.M.; van Noort, G.

DOI
10.1108/INTR-01-2021-0049

Publication date
2022

Document Version
Author accepted manuscript

Published in
Internet Research

The Algorithmic Persuasion Framework in online communication: Conceptualization and a future research agenda

Brahim Zarouali, Sophie C. Boerman, Hilde A.M. Voorveld & Guda van Noort

Amsterdam School of Communication Research

University of Amsterdam

Please cite the original published article in Internet Research:

Zarouali, B., Boerman, S.C., Voorveld, H.A.M. and van Noort, G. (2022), "The algorithmic persuasion framework in online communication: conceptualization and a future research agenda", *Internet Research*, Vol. ahead-of-print No. ahead-of-print. https://doi.org/10.1108/INTR-01-2021-0049

Structured abstract

Purpose – In our increasingly data-driven media landscape, algorithms play an important role in the consumption of online content. This article presents a novel conceptual framework to investigate algorithm-mediated persuasion processes and their effects in online communication.

Design/methodology/approach – The article introduces a comprehensive and dynamic framework that focuses on the role of algorithms in persuasive communication: the *algorithmic persuasion framework* (APF).

Findings – The APF consists of five conceptual components: *input, algorithm, persuasion attempt, persuasion process, and persuasion effects*. In short, it addresses how data variables are inputs for different algorithmic techniques and algorithmic objectives, which influences the manifestations of algorithm-mediated persuasion attempts, informing how such attempts are processed and their intended and unintended persuasive effects.

Originality/value – The article guides future research by addressing key elements in the framework and the relationship between them, proposing a research agenda (with specific research questions and hypotheses) and discussing methodological challenges and opportunities for the future investigation of the framework.

Keywords*:* Algorithm, input, persuasion attempt, persuasion process, persuasive effects

Introduction

In modern society, algorithms are often used in all sorts of information systems. An algorithm is a set of step-by-step instructions computers are programmed to follow to accomplish certain tasks (Zhong, 2021). They can be used in many contexts and for various purposes; for example, banks may use them to approve or decline mortgage applications, government agencies may use them to allocate funding requests, and criminal justice systems may use them to evaluate who might be eligible for early release (Araujo *et al.*, 2020; Fry, 2019). In addition to these examples, algorithms are also often used in data-driven media landscapes to automatically and dynamically modify the content users see. More precisely, they are used to filter enormous amounts of content and present personalized information, services, and advertisements to online users (Ricci *et al.*, 2015).

Algorithms are used across various online platforms (e.g., social media, websites, streaming services, apps, forums, etc.) and in various communication contexts, such as in commercial, political, health, or news environments (Agudo and Matute, 2021; Zarouali *et al.*, 2021). For instance, algorithms are used to show users recommended products on e-commerce websites to increase the likelihood of purchasing a product (commercial); to show a recommended workout schedule on a fitness app to increase the likelihood of getting fit (health); to show a sponsored political post on social media to increase the likelihood of voting for a party (political); or show a recommended news article to increase the likelihood of spending more time on a news platform to increase revenues (news). As these illustrations show, algorithms play an important role in how users navigate, select, and consume information in online settings. This means that users are more influenced than ever by (partially) personalized and distinct streams of online content, which are largely based on their own (and/or those of similar others) past choices and preferences (Beer, 2017; Bucher, 2018; Ricci, 2015; Zarouali *et al.*, 2021). As a result, algorithms are transforming online platforms into codified environments that expose users to the content that is likely to be most persuasive to them (Yeung, 2017).

Given the degree to which algorithms have changed the nature and type of communication people encounter online, concerns have arisen about the persuasive impact of such algorithms. These concerns include a fragmented public sphere, a higher likelihood of (voter or consumer) manipulation, an increase in attitudinal polarization, more privacy infringements, an increase in user surveillance, and a loss of user autonomy (e.g., Cho *et al.*, 2020; Susser, 2019; Tufekci, 2015; Zarouali *et al.*, 2020a). Thus, algorithmic technologies have given rise to new societal questions, about which scholars require a solid base of empirical knowledge before they can make valuable contributions to public debate. Therefore, we need answers based on insights regarding the ethical and societal consequences of persuasive algorithms in our digital society. These insights are crucial for guiding us in the direction of an open society that benefits from algorithmic developments while avoiding its potential risks. And with algorithms increasingly being used in more domains in our daily life, this has become more important than ever.

Despite the importance of studying the persuasive influence of algorithms, little is known about the multifaceted role of algorithms in how users interpret and interact with online information and communication, as well as how and when this communication subsequently influences their decisions and evaluations. While many scholars have addressed this emerging topic (e.g., Cho *et al.*, 2020; Gunaratne *et al.*, 2018; Kim *et al.*, 2019; Zarouali *et al.*, 2020a), there is a lack of overarching theoretical frameworks that would allow scholars to derive specific research questions and hypotheses. We address this gap by presenting a comprehensive framework of the role of algorithms in persuasive communication. We introduce the concept of algorithmic persuasion *and* provide a substantiated conceptual framework for the research of this phenomenon, the *algorithmic persuasion framework*

(APF), building on the *personalization process model* by Vesanen and Raulas (2006) and the *advertising process circle* by Rodgers and Thorson (2019). In particular, the goals of this paper are threefold. First, we define and conceptualize algorithmic persuasion. Second, we propose a framework that integrates different dynamics, i.e., the data input, the algorithms, the persuasion attempts, the persuasion processes, and persuasion effects of algorithm-mediated online communication. Third, we develop a future research agenda based on the insights derived from this framework.

Ultimately, this paper aims to make contributions on different levels. From a theoretical point-of-view, the APF is a novel framework that can advance the study of algorithmic persuasion in the field of communication science. This framework may lead to insights that were previously unnoticed, therefore extending our knowledge; or insights that challenge previous findings, possibly adjusting our knowledge; or insights that reinforce previous findings, hereby creating a more robust knowledge basis. In addition, studies based on this framework might contribute relevant and timely insights to the prominent societal debate on the (persuasive) *role* and *impact* of algorithms in our digital and data-driven society. Furthermore, the framework offers practitioners a useful framework for developing campaigns driven by algorithms. It provides an overview of all the key factors (e.g., the data, algorithm, and message) that need to be considered before launching communication campaigns. The insights provided by the APF can therefore lead to more effective and impactful communication efforts in practice.

Defining Algorithmic Persuasion in Online Communication

Algorithms have transformed online environments into persuasive architectures that influence the online choices of media users through processes that are subtle and unobtrusive. In other words, algorithms have become very persistent and pervasive when it comes to codified persuasion in online environments (Rose and MacGregor, 2021). This development warrants a clear and unambiguous conceptual framework to allow understanding of the persuasive effects of exposure to algorithmically tailored communications in online settings. To this end, we first introduce the concept of *algorithmic persuasion*. We define algorithmic persuasion *as any deliberate attempt by a persuader to influence the beliefs, attitudes, and behaviors of people through online communication that is mediated by algorithms*. We discuss this concept by decomposing the definition. We will discuss each element in the definition from the perspective of relevant and recent communication literature. This shows that our definition flows directly from existing work that defines persuasion or advertising (e.g., Dahlen and Rosengren, 2016), and that algorithmic persuasion bears several similarities to definitions of "traditional" or non-algorithmic persuasion. That said, we pay more attention to the unique elements of the construct (in comparison to traditional persuasion), its algorithm-mediated and circular nature.

The term *deliberate attempt* in reference to algorithmic persuasion means that the attempt is purposefully initiated by the persuader. It is important to note that the attempt to persuade does not have to be recognized by the receiver; the focus here is on the intent of the persuader. A *persuader* can be a brand, organization, or person (Dahlen and Rosengren, 2016). To illustrate, algorithmic persuasion can be employed by commercial brands, perhaps through the use of algorithms that present to consumers relevant product offerings; by news organizations in, for example, the presentation of algorithm-curated news content to its readers to increase readership; by big tech companies, who might use (filtering) algorithms to present personalized search results that increase user engagement with the platform(s); or by political candidates, who could use algorithms to target citizens with campaign messages to generate votes. These are only examples of the many types of persuaders and the almost unlimited types of persuasive influences they might generate.

Influence refers to a change in people's beliefs, attitudes, and behaviors (what people think, feel, and do) after exposure to algorithm-mediated communications (Kim, 2014). This influence can vary from very weak to very strong and can happen on a conscious or unconscious level, occurring with or without consumers' awareness. Furthermore, the influence can be short term or long term and intended or unintended (e.g., Dahlen and Rosengren, 2016). For instance, algorithms that recommend certain relevant products to consumers might lead to immediate, short-term effects (i.e., buying behavior); but alternatively, algorithms that recommend news articles that mainly fit a person's own ideological views on society (i.e., selective news exposure) might influence a person's opinions to become more extreme (and polarized) over the long term (Cho *et al.*, 2020; Pariser, 2011; Sunstein, 2009).

Online communication refers to the transmission of algorithm-mediated communication from senders to receivers in online environments; in other words, algorithmic persuasion is always an online communicative activity driven by algorithms. This communication can take place on a plethora of media devices (e.g., desktop and mobile), spanning different mediated contexts (e.g., e-commerce websites, social media platforms, search engines, and health apps). Finally, *algorithm-mediated* indicates that the online communication must be mediated by algorithms that automatically decide which content to select and present to which users based on a large corpus of input data (i.e., people's digital footprints). This last element is further explained and theorized in the next section in which we introduce the APF.

The Algorithmic Persuasion Framework (APF)

We envision the APF as a dynamic process. Our framework (see Figure 1) consists of five components that are central to algorithmic persuasion: the *input*, *algorithm*, *persuasion* *attempt, persuasion process, and persuasion effects*. These five components are presented in a circular way (i.e., all components form a chain of events in a cyclic or circular process) so that the APF accounts for feedback loops. These feedback loops are pathways by which some of the output of a particular “persuasion attempt” becomes new input in a circular process that ultimately reinforces or alters existing algorithmic systems. Feedback loops (often) manifest themselves between the persuasion effects and data input: how a user reacts to an algorithm-mediated message (the effects) will generate new data that serves as input for a new algorithm-based persuasion attempt (Julier, 2017; Vasudevan, 2020). Therefore, persuasion effects and data input can be seen as both a cause and an outcome. The circular nature of algorithmic persuasion is the key factor that differentiates algorithmic persuasion from traditional persuasion.

The different building blocks of the APF are strongly based on existing theories. It combines building blocks that originate from theories on traditional persuasion, with building blocks that originate from more modern theorizing about algorithms and personalization. The circular dynamic of our framework originates from Vesanen and Raulas’ (2006) *personalization process model*, which conceptualizes personalization as a process, showing how phases of personalization (e.g., customer data, customer profile, marketing output, etc.) are linked. Interestingly, this model posits that all these phases are related to each other in the form of a dynamic loop. Since we argue that algorithmic persuasion consists of different phases or building blocks that are interrelated, the *personalization process model* provides a solid base for the process nature and feedback loops in our framework. In addition, some of the building blocks of our model flow directly from the personalization phases in this model (e.g., data and the persuasion attempt).

Our framework is further complemented with building blocks originating from theories on traditional or non-algorithmically mediated persuasion. The framework has similarities to the *advertising process circle*, described by Rodgers and Thorson (2019). This model illustrates a cohesive set of processes related to digital advertising (and persuasive communication in general) that encompass most of the work from scholars (e.g., audiences, devices, media channels, contexts, message sources, etc.). It synthesizes all important aspects of various occurrences that have been studied in this field and provides a useful way to organize existing theories and research. It lists some factors that we also list in our framework (e.g., intended and unintended effects) but most importantly signals that there is a cohesive set of processes related to advertising and persuasive communication. In the next section, we discuss all the separate components, including their origins, and our methodological considerations. The main research questions associated with each component are summarized in Table I.

[INSERT FIGURE 1 ABOUT HERE]

The input: data

The first component of the APF is referred to as *input* which involves all data that is used in algorithmic persuasion. This component strongly corresponds with the "customer data" element mentioned in the *personalization process model* (Vesanen and Raulas, 2006). *Data* are vital for algorithmic persuasion. Before algorithm-mediated communication can be provided to online users, data must be collected and readied for the algorithm (Gillespie, 2014; Kitchin, 2014; Tafesse, 2020). Generally, a distinction is made between first-, second-, and third-party data (Schneider *et al.*, 2017; Yun *et al.*, 2020).

First-party data are data that are owned or collected by the sender, the source of the persuasion attempt itself (Canapa, 2020). These data can be collected by means of cookies on their own website, but data about online purchases, data entered when the receiver becomes a member of or donator to a political party, or data that are disclosed during registration of a device such as a fitness tracker can also be seen as first-party data. Second-party data are data that can be used for algorithmic persuasion because they are owned by a collaborating party (Sinclair, 2020). For example, when buying media space in a programmatic way (i.e., buying media space in an automated way based on real-time bidding), persuaders can make use of data owned by Google if they use the demand- or supply-side platforms provided by Google. Third-party data are collected by companies that are not directly involved in the primary process (Bresciani *et al.*, 2021). Persuaders can purchase such data from data brokers that specialize in collecting and combining data. Another important distinction is between *explicit* and *implicit* data (Buchmann, 2014; Taylor *et al.*, 2009). The former refers to data wittingly disclosed by users in online environments, whereas the latter refers to the compilation of and/or inferences from data about users collected without their awareness (e.g., surfing history, preferences, IP-address).

In sum, it can be said that online users give up their personal data either wittingly or unwittingly, in various online “roles”: as purchasers, subscribers, registrants, donors, members, etc. Therefore, understanding the persuasive potential of algorithm-mediated communication should first start with a critical examination of the (big) data that serve as sources of input.

The algorithm: techniques, persuader objectives, and biases

After collection and storage, the data are transformed (e.g., data cleaning, filtering, selecting, etc.) so that algorithms can act automatically; i.e., the data are converted into an “algorithm-friendly” format (Gillespie, 2014), which brings us to the next factor in the model: the algorithm. Although the algorithm is not mentioned as a component in Vesanen and Raulas’ (2006) model, it is implied by what they called processing (e.g., profiling and segmenting) in their model. Algorithms can be understood as encoded procedures for transforming input data into a desired output, based on specified calculations (Gillespie, 2014). It is important to address that “the” algorithm does not exist: there is no algorithm that provides one-size-fits-all solution for all persuaders (Kalpokas, 2019).

There are many different algorithms, each with their own goals, idiosyncrasies advantages, and drawbacks (Fry, 2019). But broadly speaking, algorithms can be divided into two categories: (simple) rule-based algorithms and (complex) machine learning algorithms (Bucher, 2018; Edwards and Veale, 2017; Kalpokas, 2019). These two categories, which have different applications, can guide persuaders in building algorithmic systems that meet their specific objectives (see next paragraph), for instance, the use of natural language processing in chatbots and virtual assistants (Araujo *et al.*, 2019), the use of recommender systems to tailor online content (Ricci, 2015), or the use of conversational search in smart devices with voice interfaces (Jung *et al.*, 2019). On a more specific level, we identified four specific *algorithmic techniques* that are widely used: prioritization (making an ordered list), classification (picking a category), association (finding links), and filtering (isolating what’s important; Diakopoulos, 2016; Fry, 2019). Most algorithms are developed based on a combination of the above techniques, depending on the specific objectives or goals of the persuader, which brings us to a second important point.

As addressed, how an algorithm is built will depend heavily on the *objective* of the persuader. In persuasive communication, a traditional distinction is made between cognitive, affective, and behavioral influence as persuasive aims (e.g., Eisend and Tarrahi, 2016; Perloff, 2017; Vakratsas and Ambler, 1999). With one or more of these objectives in mind, algorithms can be set to work to achieve their aim. For instance, persuasion objectives related to cognitive outcomes could include increasing memory of the persuasion attempt or evoking certain thoughts about the persuader (i.e., by focusing on views, impressions, etc.). Aims related to affective outcomes might include encouraging positive attitudes towards the persuader or promoting liking of the persuasive attempt or positive evaluations of the provided content. Aims related to behavioral outcomes include ensuring continued use of a media platform (i.e., time spent on platform), increasing interaction with content offered by the platform (e.g., liking, sharing, and commenting), and prompting users to click on items, buy products online or offline, donate money for a political party or candidate, etc. Thus, using algorithms to provide personalized media content and services allows persuaders to achieve these cognitive, affective, and behavioral aims and increase user engagement (van Dijck *et al.*, 2018; Rose and MacGregor, 2021). Therefore, we argue that algorithmic persuasion is heavily influenced by persuaders' choices of what should be prioritized and for what purpose or objective.

A final note in this section relates to algorithmic bias. The use of personal data and algorithms for the purpose of persuasion comes with ethical and social considerations, such as issues with privacy (Acquisti *et al.*, 2015), information asymmetry (Mittelstadt *et al.*, 2016), covert manipulation (Koene *et al.*, 2015), and bias (Bozdag, 2013). Biases emerge because developers unconsciously program their biases (e.g., prejudices and stereotypes) into algorithms and/or because machine learning models have been trained on flawed and biased datasets (Seeber *et al.*, 2020). In addition, algorithms are created for purposes that are often far from neutral: they are made to persuade, to increase value and capital; to nudge behavior and change people's preferences; and to identify, sort, and classify people (Kitchin, 2017). Biased algorithms are a major problem of algorithmic persuasion (Fry, 2019; Willson, 2017; Zarouali *et al.*, 2021). To illustrate, a recent study found evidence of gender bias in Google's image search results (i.e., systematic underrepresentation of women), meaning that the search engine returns biased imagery to its users (Kay *et al.*, 2015). As this example highlights, algorithms can exert a persuasive influence on people's views and beliefs by presenting them biased information and communication. This is an important element that could lead to undesired persuasion effects (see further).

The persuasion attempt: context, nature, medium and modality

In the next phase, inputs (i.e., data) and algorithms give shape to the actual algorithm-mediated persuasion attempt. Elements of this phase are similar to the elements called marketing output and delivery in the *personalization process model* (Vesanen and Raulas, 2006). How algorithmic communication manifests itself to a receiver depends on the inputs and the algorithmic codes. This means that the receiver may encounter and observe algorithmic persuasion in very different ways and across diverse communication contexts. Again, the deliberative attempt of the persuader does not have to be recognized as such by the receiver; here we solely focus on how the receiver may encounter algorithmic persuasion, whether this is consciously observed and recognized as an attempt or not. We argue that to fully understand algorithmic persuasion, it is important to consider the full spectrum of how persuasive communication manifests itself, including the communication *contexts*, the *nature* of the communication, and the *medium* and *modality* through which it is expressed.

First, communication research is built around communication contexts (e.g., advertising, corporate, and health communication) that are also relevant for algorithmic persuasion attempts. Algorithmic persuasive attempts occur in different contexts, which mainly differ because the persuader can be anyone or any type of organization, including brands, governmental bodies, and non-profit organizations (Zarouali *et al.*, 2021). Thus, algorithmic persuasive attempts occur in commercial and non-commercial contexts, including health, e-commerce or advertising, news, and political communication (van Dijck *et al.*, 2018; Rose and MacGregor, 2021). This means that algorithmic communication is not limited to the commercial content (e.g., advertising) but includes public relations, organization and strategic communication, health information campaigning, political communication, and so on, essentially, any organized communication intended to have a certain impact. Studies have demonstrated that communication contexts moderate the impacts of persuasive communication. Previous studies specifically related to algorithmic persuasion, for example, demonstrated that AI influences people differently in different communication contexts (Agudo and Matute, 2021). In addition, these contexts may moderate the persuasion outcomes of data-driven communication (Bol *et al.*, 2018). In a similar vein, the communication context may play an important role in the outcomes of algorithmic persuasion.

Second, algorithm-driven persuasive attempts can also differ in nature: they can present as *paid* content from some persuader (in line with Richards and Curran, 2002), often called "sponsored" on social media platforms, or as *organic* in nature. The development that persuasion no longer takes the primary form of paid content is extensively described in the conceptual work of Dahlen and Rosengren (2016). Similar distinctions, for example, between paid, owned, and earned content originating from marketers, have also been described in the academic literature (e.g., Goodall, 2009; Stephen and Galak, 2012).

A paid algorithmic persuasion attempt usually takes the form of personalized advertising (or "sponsored content") that appears with the same form and qualities as a platform's original content and is usually meant to favorably influence users' perceptions of the third-party sponsor (Sonderman and Tran, 2013). Organic content refers to content that is contributed to a site by the users themselves (e.g., user-generated posts on social media, all websites on search engines, and all videos on video-sharing platforms; Proffitt, 2011). Big tech platforms (e.g., Facebook, YouTube, and Google) and other companies (e.g., news organizations) have their own proprietary algorithms that determine the visibility of organic content to users (not everyone sees the same organic content), which can have a significant persuasive influence in shaping people's views and decisions (Epstein and Robertson, 2015; Tufekci, 2014). Thus, algorithmic persuasion is not limited to "paid persuasion" (i.e., entities that pay to use algorithms to persuade individuals about specific issues) but can also be the result of the algorithmic determinations of big platforms on what organic content should be prioritized (i.e., entities whose business model is based on using algorithms to direct and commodify the attention of individuals; Helberger, 2019).

Third, algorithm-mediated persuasive communication can be distributed via diverse online channels (e.g., search websites, social media platforms, e-commerce sites, streaming services, news websites, and so on) and via a variety of devices, including smart TVs, smartphones, tablets, laptops, and smart speakers. In traditional persuasion models like that of Laswell (1948) or McGuire (1989), this is referred to as the medium, the instrument that brings the advertising message to its audience (Rodgers and Thorson, 2012). For a long time, scholars have argued that the medium can be as important as the communication itself (McLuhan, 1964). That is, algorithm-mediated content is profoundly shaped by the characteristics of a given medium or user interface (Sundar, 2020; Sundar *et al.*, 2015). The affordances of a given medium will influence how people process and evaluate algorithm-mediated communication, based on objective characteristics (e.g., pacing, interactivity, ephemerality) as well as the way different media are experienced (e.g., Bronner and Neijens, 2006; Voorveld *et al.*, 2018). For example, a recommended news item will not appear in the same way across different media because the algorithms embedded in each medium will shape the content based on its unique, distinct rules (Sundar, 2020). Thus, the impact of algorithm-mediated communication cannot be studied in isolation from the medium in which it is embedded.

Fourth, given the variation in modes of delivery, the modality of algorithm-mediated communication also varies. Such communication can, for example, be presented in either (audio)visual form (e.g., a recommended post, a recommended video, etc.) or auditory form (e.g., voice assistant-recommended content). Prior work has clearly demonstrated that modality influences the persuasiveness of communication for offline communication modes including audio, video, and written messages (Chaiken and Eagly, 1976). Theories like the MAIN model (Sundar, 2008), and the cognitive theory of multimedia learning (Mayer, 2014) acknowledged the importance of modality in shaping communication effects. Recently, studies have been conducted that demonstrated the modality of algorithm-mediated communication (e.g., voice vs. text) plays a role in its persuasive impact on recipients (Cho, 2019; Cho *et al.*, 2019; Voorveld and Araujo, 2020). This implies that modality should be considered in the investigation of outcomes of algorithm-mediated communication.

The persuasion process: five underlying mechanisms

Having discussed how algorithm-mediated communication manifests itself in the form of persuasive attempts, we now discuss how these attempts are processed by recipients. We discuss five *underlying processes* through which algorithms act as persuasive tools, altering the cognitions, attitudes, or behaviors of online users: *relevance*, *reduction*, *social norms, automation,* and *reinforcement*. This list of mechanisms is not exhaustive, but it provides a helpful basis for understanding some of the important underlying dynamics involved in algorithmic persuasion. The included processes were derived from existing empirical research on the media effects of algorithm-mediated content. In the next paragraphs, we elaborate on the theoretical underpinnings of how these principles might explain the ways algorithm-mediated content persuades people.

A first process of persuasion is related to the concept of *relevance*. Essentially, what algorithms do is select and present – based on a data-driven approach – content that aligns well with the interests and preferences of the recipients (Ricci, 2015). This tailoring centers around the idea of increasing the personal relevance of presented content, which favorably shapes attitudes and behaviors with regards to the presented items (Boerman *et al.*, 2017; e.g., De Keyzer *et al.*, 2015). In other words, algorithms can be seen as tailoring technologies that may achieve persuasion by increasing the personal relevance of the recommended content.

A second process of persuasion focuses on the ability of algorithms to *reduce* a very large corpus of content into a smaller consideration set to avoid choice overload (e.g., search results, news feed, etc.; Rader and Gray, 2015; Ricci, 2015). Psychological and economic theories suggest that people perform mental cost-benefits analysis when using technology, seeking to minimize costs and maximize benefits (Fogg, 2003; Kaptein *et al.*, 2012). By reducing the media choice to a limited number of (relevant) content options, algorithms can increase the perceived favorability of those options (Fogg, 2003).

A third process involves algorithms' ability to persuade people by showing them the preferences and behaviors of others, which sets up *social norms* (Cialdini and Trost, 1998; Fogg, 2003). Algorithms recommend content to people based on what (similar) others like (i.e., collaborative filtering; Ricci, 2015). These items are often supplemented with signals of social norms, such as other customers' ratings, friends' likes, etc. Studies have shown that these social cues increase the likelihood of eliciting more favorable attitudes toward specific content and generating behavioral change (Bakshy *et al.*, 2012; Zarouali *et al.*, 2020b). Therefore, algorithmic persuasion can also be the result of people's susceptibility to social norms.

As a fourth process, algorithmic recommendations can be persuasive as a result of *automation bias*, which refers to people attributing greater trust in machines and their recommendations compared to other sources of recommendation (Sundar, 2020). Automated decisions made by algorithms can be perceived as more objective and rational than decisions made by humans (Clerwall, 2014; Seeger and Heinzl, 2018) and as a result, lead to more favorable responses to content provided by algorithms (e.g., Araujo *et al.*, 2020; Graefe *et al.*, 2018). Therefore, the automation bias mechanism could serve an explanation of how users evaluate algorithm-driven communication.

A fifth persuasion process relates to the potential of using algorithms to *reinforce* people's pre-existing attitudes and views (Ohme, 2021; Sunstein, 2009). Because algorithms are often used to selectively show people content they would like to see, persuasion can occur through a process of attitude-reinforcement. In this respect, research has shown that algorithmic content that reinforces people's prior attitudes leads to more clicks on the content, as well as an increase in likelihood of accepting or agreeing with the information included in the content (Beam, 2014; Cho *et al.*, 2020; Dylko *et al.*, 2017).

The persuasive effects: intended and unintended effects

The last component of our framework involves the persuasive effects of algorithm-curated communication after exposure to and processing of an algorithmic persuasion attempt (i.e., how users respond to it). Given the circular dynamics of the framework, it assumes that users' responses to algorithm-recommended communication also function as (data) input for new communication and selection processes (Liu-Thompkins *et al.*, 2020). This is also in line with Vesanen's (2007) model on personalization. For instance, clicks on a recommended news article can serve as useful input of a person's interest in the particular issue or topic. In addition, advertisers can use people's responses to a campaign as input for new campaign efforts.

Traditional communication models such as Shannon and Weaver's Model of Communication (McQuail, 2008) usually include some kind of interfering variable ("noise"), indicating that communication may not always have the anticipated outcome. Following traditional communication theories, the *advertising process circle* (Thorson and Rodgers, 2019), seminal research into persuasion and advertising effects (Buijzen and Valkenburg, 2003; Cho and Salmon, 2007), and models addressing persuasion and technology (e.g., Berdichevsky and Neuenschwander, 1999; Reaves, 1995), the APF incorporates the notion that algorithmic persuasion can have both *intended* effects and *unintended* effects. Intended effects are those effects that are desired by the persuader exposing people to the algorithmically recommended content. Unintended effects are secondary, usually undesired effects of exposure to algorithm-mediated content (Buijzen and Valkenburg, 2003; Reaves, 1995). Both intended and unintended effects can be ethical or unethical, and unintended effects differ with respect to whether they are predictable or not (Berdichevsky and Neuenschwander, 1999).

There are several important unintended effects that apply to algorithmic persuasion in general. The use of algorithms to filter, order, and personalize messages is unavoidably linked to covert behavior manipulation (Koene *et al.*, 2015). What information is shown to people changes the options they have, and thus their behavior, and this often occurs without the user's awareness. Additionally, collecting the data required for algorithmic persuasion comes with significant privacy issues (Acquisti *et al.*, 2015) and may result in information asymmetry: an imbalance in knowledge and decision-making power favoring data processors over the user (Mittelstadt *et al.*, 2016). Because typical users do not know how algorithms work and thus cannot control, monitor or correct them, this significantly reduces their agency, which is an important undesired consequence of algorithmic persuasion. Furthermore, biases in algorithms can have the unintended effect of influencing people's views and beliefs, and such biases can lead to discrimination in the recommendations made or the messages shown to specific people (e.g., Ali *et al.*, 2019; Bozdag, 2013; Obermeyer *et al.*, 2019) To further illustrate the intended and unintended consequences of algorithmic persuasion more concretely, we give examples of these possible effects in different communication contexts including advertising, political communication, news, and health communication. Research in communication science is built around such contexts (Rigotti and Rocci, 2006).

When algorithmic persuasion is used in the context of marketing communication, intended effects include increased brand awareness, brand attitudes, patronage intentions, purchase intentions, sales, click through rates, and brand loyalty (Bakpayev *et al.*, 2020; e.g., Boerman *et al.*, 2017; van Esch *et al.*, 2021; Matz *et al.*, 2017; Van den Broeck *et al.*, 2019). For instance, research has shown that algorithmically tailoring online advertising based on personal information such as online behavior (Bleier and Eisenbeiss, 2015), social media likes and demographic characteristics (Aguirre *et al.*, 2015), and psychological characteristics (Matz *et al.*, 2017) increases such intended effects. Potential unintended effects of algorithm-mediated advertising include covert persuasion, where people's behavior and attitudes are influenced outside their awareness (Savage, 2019), intrusion into consumers' (data) privacy (Boerman *et al*., 2017), and the targeting of users with unfair commercial tactics to exploit their personal weaknesses and vulnerabilities (e.g., impulsive buying, confusion; Dāvida, 2020).

Algorithmic persuasion often occurs in political communication in the form of microtargeting. Political microtargeting can have the intended effects of exposing people to more relevant political information, reaching social groups that are difficult to contact, and increasing voter knowledge about individually relevant issues, thus influencing voter's attitudes and behavior (Zarouali *et al.*, 2020a; Zuiderveen Borgesius *et al.*, 2018). However, potential unintended and negative effects include manipulation of voters, algorithmic exclusion of voter groups, and a lack of algorithmic transparency (for an extensive overview, see Zuiderveen Borgesius *et al.*, 2018). In addition, big data can be used to identify citizens' "decision-making vulnerabilities," and algorithms can subsequently be used to deliver tailored messages that exploit these vulnerabilities (Zarouali *et al.*, 2020a). Finally, political microtargeting may lead to undesirable behavior from a democratic viewpoint, with voters refraining from certain behavior – such as seeking information about the stances of political parties – because they perceive they are being watched (the so-called chilling effect; Dobber *et al.*, 2019).

In the news context, algorithmic persuasion occurs in the form of recommender systems that produce personalized news content. News recommender systems have the intended effect of increasing readers' engagement on the news platform and increasing news consumption, thus positively influencing the extent to which people are informed about current affairs (DeVito, 2017; Helberger, 2019). According to some researchers, however, an unintended consequence is that consistently exposure to algorithmically-curated news leads to "filter bubbles" (Pariser, 2011) or "echo-chambers" (Sunstein, 2009). The idea behind this claim is that algorithms will selectively show people only content they would like to see. This reinforces their pre-existing attitudes and views on all sorts of issues, and as a result, they get trapped in "like-minded" information bubbles that are completely isolated from different or opposing viewpoints (Dylko *et al.*, 2017; Flaxman *et al.*, 2016). In the long term, this could contribute to increased polarization in society (Cho *et al.*, 2020; Flaxman *et al.*, 2016; Pariser, 2011). These unintended effects might not occur immediately, in the short term, but could come to the surface after longer periods of systematic exposure to algorithmically recommended content.

Algorithmic persuasion is also important in health communication. Algorithmic recommender systems support and enhance digital health programs, helping people stop smoking and engage in healthier behaviors (Cheung *et al.*, 2019; Kim *et al.*, 2019). In addition, there has been spectacular growth in online services and mobile health apps (e.g., fitness and medical apps) that use algorithms to advance the health, fitness, and physical or mental well-being of users (van Dijck *et al.*, 2018). E-health developments, such as algorithm-driven health programs and services, arguably facilitate "socio-economic inclusion and equality, quality of life, and patient empowerment through greater transparency, access to services and information and the use of social media for health" (European Commission, 2012). Thus, intended effects of algorithmic persuasion in the context of health communication include enabling people to monitor their own health behavior, positively influencing healthy behavior, and to connect people with the same health issues. Unintended, and potentially negative outcomes of the growth of and reliance on algorithmically driven health apps and services is the creation of new digital inequalities (Bol *et al.*, 2018); the targeting (and influencing) of people based on sensitive, health-related patient data; and problematic collection, use, and commodification of sensitive personal data about health (van Dijck *et al.*, 2018; Liang *et al.*, 2021).

[INSERT TABLE I ABOUT HERE]

Methodological considerations

In this section, we discuss the methodological approaches we used to investigate different components of the APF. First, certain hurdles made researching the first two components, input data and algorithms, difficult. The mechanics of many real-life algorithms and the data that feed them are largely hidden and not open to empirical research, turning them into impenetrable black boxes (Beer, 2017; Bucher, 2018; van Dijck *et al.*, 2018; Lomborg and Kapsch, 2020). For many companies, these algorithms serve as a main competitive advantage, which they are reluctant to expose to others (Thompson *et al.*, 2021). Even without access to the specific algorithms, scholars have proposed interesting methodological approaches, such as algorithmic auditing and reverse engineering (Bodo *et al.*, 2017; Kitchin, 2017; Liang *et al.*, 2021). However, these methods come with considerable technical and practical challenges for social science scholars. Therefore, qualitative research is likely the most feasible option for future research. Such qualitative research would involve scholars setting up interviews with algorithmic developers and persuaders (if they can find enough who are willing, which might be the biggest obstacle) to uncover the nature of the data, the logic of the algorithms, and the story behind the persuasion objectives and goals.

Second, we argue that algorithmic persuasion attempts, processes, and effects can be investigated based on at least three methodological approaches within the field of social sciences: *scenario-based studies; browser plug-in studies;* and *mock environment studies*.

The first method, a scenario-based experiment, is the most conventional and an often-used way to gain insights into people's responses to and perceptions of tailored persuasive communication without collecting large amounts of personal data and creating personalized messages for all participants (Boerman *et al.*, 2017). This approach usually includes some kind of scenario-based experimental procedure. Based on some initial questions (i.e., input data), participants are exposed to a persuasion attempt (e.g., a news article, an ad, a health message, etc.) that is tailored to their initial answers (e.g., Beam, 2014; De Keyzer *et al.*, 2015; Sutanto *et al.*, 2013), and the researchers subsequently measure the processes and effects. The aim of this method is to mimic the functioning of an algorithm by showing participants content that includes some part of their disclosed personal preferences. Although this is an easy and straightforward approach, it comes with some downsides: these studies do not include a real recommendation algorithm and tend to be low in ecological validity.

A second method is the use of a custom-built browser plug-in designed to collect participants' online interactions and behaviors (e.g., Bodo *et al.*, 2017; Haim and Nienierza, 2019; Strycharz *et al.*, 2019). This approach has also been referred to as "data donation" (Krafft *et al.*, 2019). A browser plug-in allows researchers to unobtrusively track which pieces of algorithm-mediated information participants are exposed to on their own devices. In addition, the plug-in also allows the collection of behaviors as a result of exposure to algorithm-mediated communication, such as comments, likes, shares, clicks, follow upsearches, etc. With these data, quantitative statistical analyses are possible across different users, browsers, and settings (Haim and Nienierza, 2019), allowing quantification of micro-level algorithmic persuasion effects. The main challenges related to this approach are legal and ethical in nature: the privacy of people should be taken as a serious concern, and there are data restrictions imposed by the platforms and organizations that provide gatekeeping services.

Third, scholars could build a simplified version of an algorithm-based (mock) environment, which fits their own particular research context. An example of this approach can be found in research by Loecherbach and Trilling (2020), who developed their own news recommender system to test the media effects of personalized news on users. Another illustration is a recent study by Zarouali *et al.* (2020a), for which they developed a mock social networking site to expose users to political ads driven by algorithms, allowing the researchers to determine the underlying mechanisms and persuasion effects of political microtargeting. These experimental settings represented fully controlled environments (i.e., beneficial for ecological validity) in which real algorithms were used to test persuasion-related processes and effects. In addition, this approach also provides great opportunities to make such settings open-source to increase transparency, collaboration, and replicability of future algorithmic persuasion research.

General discussion

In recent years, many computational and automatization processes have found their way into different digital media and technologies, bringing forth an increase in algorithm-mediated persuasive communication. While many scholars have paid attention to these technological trends, there is a lack of theoretical frameworks that allow scholars to derive specific research questions and hypotheses. To advance this emerging research field, we introduced the *algorithmic persuasion framework* (APF; see Figure 1). This integrative, dynamic theoretical framework consists of five components that are related in a circular way and guides scholars through the process of how algorithm-mediated persuasion comes about, as well as how it exerts persuasive influence on online users. To further advance this field, we proposed a future research agenda, based on this framework, that addresses both theoretical and methodological challenges.

In terms of the future research agenda, the APF offers an integrative overview within which to examine the different components of algorithmic persuasion. Potential future research questions are summarized in Table I. Most of these questions are intended to guide contributions that are not only disciplinary but also cross disciplinary (Turel *et al.*, 2019). The starting point of the framework is that the effects of algorithm-recommended communication cannot be understood in isolation from the initial *data* (implicit and explicit data / first-, second-, and third-party data) that served as input, nor from the *algorithmic techniques* (prioritization, association, classification, and filtering) that were developed and deployed to achieve certain predefined persuasion objectives (e.g., increase time spent on platform, increase ad revenues, increase user engagement, etc.). Thus, this framework posits that *any* algorithmic persuasion attempt in future research needs to be understood within the context of a wider technological assemblage and decision-making processes on the part of the persuader (i.e., persuasion objectives).

Having addressed input and algorithm, we used the APF to provide an overview of how these elements translate into actual persuasion attempts. A future research agenda can be formulated based on the different manifestations of algorithm-mediated communication. We argued that algorithm-based communication varies with respect to *context*, *nature*, *medium* and *modality* variables and that any manifestation of algorithmic persuasion will be (theoretically) shaped by these factors. This suggests that hypotheses and research questions focused on the interaction of these factors and algorithm-mediated communication are needed to build a solid understanding of algorithmic persuasion in different mediated contexts.

In addition to the persuasive attempts, the APF also incorporates the *underlying processes* that might be at play when online users are exposed to an algorithmic persuasive event (relevance, reduction, social norm, automation, and reinforcement). Future research should investigate to what extent these underlying mechanisms are relevant variables when users process persuasive information recommended by algorithms. In addition, as already mentioned, the list of persuasion processes should be taken as provisional rather than exhaustive. Therefore, we point to the (possible) existence of underlying processes other than the ones discussed in this article, which could be pursued in future research.

Finally, the APF proposes that future research should make a conceptual distinction between the *intended* and *unintended persuasive effects* of algorithm-mediated communication. This distinction is crucial to understanding the desired persuasive impact of algorithms while, at the same time, identifying unanticipated effects that are incidental or even accidental. This will allow a much-needed normative exploration and discussion of the ethics of persuasive algorithms in online environments.

The APF not only guides future research but might also fuel societal debate on the (persuasive) *role* and *impact* of algorithms in our digital and data-driven society and offers practitioners a useful framework for developing communication campaigns driven by algorithms. It provides an overview of all the factors (e.g., the data, algorithm, and message) that need to be considered before launching communication campaigns. Moreover, the framework might help in determining the right metrics for the evaluation of campaign results. Currently, in automated and digital communication contexts, short-term consumer responses such as click rates, views, comments, likes, and shares are the primary measures used (van Noort *et al.*, 2020); however, a design for a campaign based on the current framework might clearly indicate that a different set of measures are more suitable: the determined objectives of the algorithm, which can be affective, cognitive, or behavioral in nature, should guide the right evaluation metrics. Also, the framework might guide ethical considerations and discussions. More concretely, in developing AI models for persuasion in accordance with AI principles (e.g., transparency, responsibility; e.g., Jobin *et al.*, 2019) and ethical codes of conduct, the algorithmic techniques adopted and biases that might occur should be considered in the process of developing the algorithmic persuasion. Although ethical principles can be difficult to translate into action (Schiff *et al.*, 2020), the academic literature does provide clear cut interventions to stimulate the creation of ethical and responsible algorithms (e.g., Morley *et al.*, 2020). In sum, the APF and insights derived from research based on the APF, could lead to more ethical and impactful communication efforts in practice and to societal debates about ethical AI in persuasion.

Conclusion

The future of media lies in algorithmic systems that leverage the increasing availability of big data and computational infrastructures. Therefore, there is a stronger need than ever before for an integrative framework that identifies how to investigate media effects resulting from exposure to persuasive, algorithm-mediated communication in online environments. This article presents and describes a novel conceptual framework called the *algorithmic persuasion framework*. It identifies the five key components that are most central to algorithmic persuasion, as well as their inter-relational dynamics, which scholars need to consider when conducting empirical research on the persuasion effects of algorithm-mediated communication. Based on the insights derived from these components, the framework offers specific research foci for future research, highlights the possibility of using different methodological approaches, and provides insights into how different key components within the framework are related to one another. Based on this contribution, we hope to advance the emerging research field of algorithmic persuasion.

References

Acquisti, A., Brandimarte, L. and Loewenstein, G. (2015), “Privacy and human behavior in the age of information”, *Science*, Vol. 347 No. 6221, pp. 509–514.

Agudo, U. and Matute, H. (2021), “The influence of algorithms on political and dating decisions”, *PLOS ONE*, Vol. 16 No. 4, pp. 1–17.

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K. and Wetzels, M. (2015), “Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness”, *Journal of Retailing*, Vol. 91 No. 1, pp. 34–49.

Ali, M., Sapiezynski, P., Bogen, M., Korolova, A., Mislove, A. and Rieke, A. (2019), “Discrimination through Optimization: How Facebook’s Ad Delivery Can Lead to Biased Outcomes”, *Proceedings of the ACM on Human-Computer Interaction*, Vol. 3 No. CSCW, pp. 1–30.

Araujo, T., Helberger, N., Kruikemeier, S. and de Vreese, C.H. (2020), “In AI we trust? Perceptions about automated decision-making by artificial intelligence”, *AI & SOCIETY*, Vol. 35 No. 3, pp. 611–623.

Araujo, T., van Zoonen, W. and ter Hoeven, C. (2019), *Automated 1-2-1 Communication*, SWOCC, Amsterdam, available at: https://www.swocc.nl/kennisbank-item/automated-1-2-1-communication/ (accessed 14 January 2020).

Bakpayev, M., Baek, T.H., van Esch, P. and Yoon, S. (2020), “Programmatic creative: AI can think but it cannot feel”, *Australasian Marketing Journal*, Vol. 30 No. 1, pp. 90–95

Bakshy, E., Eckles, D., Yan, R. and Rosenn, I. (2012), “Social influence in social advertising: evidence from field experiments”, *Proceedings of the 13th ACM Conference on Electronic Commerce*, presented at the Proceedings of the 13th ACM Conference on Electronic Commerce, ACM, pp. 146–161.

Beam, M.A. (2014), "Automating the News: How Personalized News Recommender System Design Choices Impact News Reception", *Communication Research*, Vol. 41 No. 8, pp. 1019–1041.

Beer, D. (2017), "The social power of algorithms", *Information, Communication & Society*, Vol. 20 No. 1, pp. 1–13.

Berdichevsky, D. and Neuenschwander, E. (1999), "Toward an ethics of persuasive technology", *Communications of the ACM*, Vol. 42 No. 5, pp. 51–58.

Bleier, A. and Eisenbeiss, M. (2015), "The Importance of Trust for Personalized Online Advertising", *Journal of Retailing*, Vol. 91 No. 3, pp. 390–409.

Bodo, B., Helberger, N., Irion, K., Zuiderveen Borgesius, F., Moller, J., van de Velde, B., Bol, N., van Es, B. and de Vreese, C. (2017), "Tackling the Algorithmic Control Crisis - The Technical, Legal, and Ethical Challenges of Research into Algorithmic Agents", *Yale Journal of Law and Technology*, Vol. 19, p. 133.

Boerman, S.C., Kruikemeier, S. and Borgesius, F.J.Z. (2017), "Online Behavioral Advertising: A Literature Review and Research Agenda", *Journal of Advertising*, Vol. 46 No. 3, pp. 363–376.

Bol, N., Dienlin, T., Kruikemeier, S., Sax, M., Boerman, S.C., Strycharz, J., Helberger, N. and de Vreese, C.H. (2018), "Understanding the Effects of Personalization as a Privacy Calculus: Analyzing Self-Disclosure Across Health, News, and Commerce Contexts", *Journal of Computer-Mediated Communication*, Vol. 23 No. 6, pp. 370–388.

Bozdag, E. (2013), "Bias in algorithmic filtering and personalization", *Ethics and Information Technology*, Vol. 15 No. 3, pp. 209–227.

Bresciani, S., Ferraris, A., Romano, M. and Santoro, G. (2021), "Data Management", *Digital Transformation Management for Agile Organizations: A Compass to Sail the Digital*

World, Emerald Publishing Limited, pp. 139–152.

Bronner, F. and Neijens, P. (2006), "Audience Experiences of Media Context and Embedded Advertising: A Comparison of Eight Media", *International Journal of Market Research*, Vol. 48 No. 1, pp. 81–100.

Bucher, T. (2018), *If...Then: Algorithmic Power and Politics*, Oxford University Press.

Buchmann, J. (2014), *Internet Privacy: Options for Adequate Realisation*, Springer Science & Business Media.

Buijzen, M. and Valkenburg, P.M. (2003), "The Unintended Effects of Television Advertising: A Parent-Child Survey", *Communication Research*, Vol. 30 No. 5, pp. 483–503.

Canapa, D. (2020), "Mergers, data markets and competition", *Legal Challenges of Big Data*, Edward Elgar Publishing, available at: https://www.elgaronline.com/view/edcoll/9781788976213/9781788976213.00010.xml (accessed 29 May 2021).

Chaiken, S. and Eagly, A.H. (1976), "Communication modality as a determinant of message persuasiveness and message comprehensibility", *Journal of Personality and Social Psychology*, American Psychological Association, US, Vol. 34 No. 4, pp. 605–614.

Cheung, K.L., Durusu, D., Sui, X. and de Vries, H. (2019), "How recommender systems could support and enhance computer-tailored digital health programs: A scoping review", *Digital Health*, Vol. 5, pp. 1–19.

Cho, E. (2019), "Hey Google, Can I Ask You Something in Private?", *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19*, presented at the 2019 CHI Conference, ACM Press, Glasgow, Scotland UK, pp. 1–9.

Cho, E., Molina, M.D. and Wang, J. (2019), "The Effects of Modality, Device, and Task Differences on Perceived Human Likeness of Voice-Activated Virtual Assistants", *Cyberpsychology, Behavior, and Social Networking*, Vol. 22 No. 8, pp. 515–520.

Cho, H. and Salmon, C.T. (2007), "Unintended Effects of Health Communication Campaigns", *Journal of Communication*, Vol. 57 No. 2, pp. 293–317.

Cho, J., Ahmed, S., Hilbert, M., Liu, B. and Luu, J. (2020), "Do Search Algorithms Endanger Democracy? An Experimental Investigation of Algorithm Effects on Political Polarization", *Journal of Broadcasting & Electronic Media*, Vol. 64 No. 2, pp. 150–172.

Cialdini, R.B. and Trost, M.R. (1998), "Social influence: social norms, conformity, and compliance", in Gilbert, D.T., Fiske, S.T. and Lindzey, G. (Eds.), *The Handbook of Social Psychology*, 4th edition., McGraw-Hill, pp. 151–192.

Clerwall, C. (2014), "Enter the Robot Journalist: Users' perceptions of automated content", *Journalism Practice*, Vol. 8 No. 5, pp. 519–531.

Dahlen, M. and Rosengren, S. (2016), "If Advertising Won't Die, What Will It Be? Toward a Working Definition of Advertising", *Journal of Advertising*, Vol. 45 No. 3, pp. 334–345.

Dāvida, Z. (2020), "Consumer Rights and Personalised Advertising: Risk of Exploiting Consumer Vulnerabilities", *SOCRATES. Rīga Stradiņš University Faculty of Law Electronic Scientific Journal of Law*, Vol. 1 No. 16, pp. 76–86.

De Keyzer, F., Dens, N. and Pelsmacker, P.D. (2015), "Is this for me? How Consumers Respond to Personalized Advertising on Social Network Sites", *Journal of Interactive Advertising*, Vol. 15 No. 2, pp. 124–134.

DeVito, M.A. (2017), "From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed", *Digital Journalism*, Vol. 5 No. 6, pp. 753–773.

Diakopoulos, N. (2016), "Accountability in algorithmic decision making", *Communications* *of the ACM*, Vol. 59 No. 2, pp. 56–62.

van Dijck, J., Poell, T. and de Waal, M. de. (2018), *The Platform Society*, Oxford University Press, New York.

Dobber, T., Trilling, D., Helberger, N. and de Vreese, C. (2019), "Spiraling downward: The reciprocal relation between attitude toward political behavioral targeting and privacy concerns", *New Media & Society*, Vol. 21 No. 6, pp. 1212–1231.

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M. and Aaziz, O. (2017), "The dark side of technology: An experimental investigation of the influence of customizability technology on online political selective exposure", *Computers in Human Behavior*, Vol. 73, pp. 181–190.

Edwards, L. and Veale, M. (2017), *Slave to the Algorithm? Why a "right to an Explanation" Is Probably Not the Remedy You Are Looking For*, preprint, LawArXiv, available at: https://doi.org/10.31228/osf.io/97upg (accessed 16 October 2020).

Eisend, M. and Tarrahi, F. (2016), "The Effectiveness of Advertising: A Meta-Meta-Analysis of Advertising Inputs and Outcomes", *Journal of Advertising*, Vol. 45 No. 4, pp. 519–531.

Epstein, R. and Robertson, R.E. (2015), "The search engine manipulation effect (SEME) and its possible impact on the outcomes of elections", *Proceedings of the National Academy of Sciences*, Vol. 112 No. 33, pp. 4512–4521.

van Esch, P., Cui, Y. and Jain, S.P. (2021), "Stimulating or Intimidating: The Effect of AI-Enabled In-Store Communication on Consumer Patronage Likelihood", *Journal of Advertising*, Vol. 50 No. 1, pp. 63–80.

European Commission. (2012), *EHealth Action Plan 2012-2020 - Innovative Healthcare for the 21st Century*, European Commission, Brussels, available at: https://ec.europa.eu/health/sites/health/files/ehealth/docs/com_2012_736_en.pdf (accessed 15 May 2020).

Flaxman, S., Goel, S. and Rao, J.M. (2016), “Filter Bubbles, Echo Chambers, and Online News Consumption”, *Public Opinion Quarterly*, Vol. 80 No. S1, pp. 298–320.

Fogg, B.J. (2003), *Persuasive Technology: Using Computers to Change What We Think and Do*, Morgan Kaufmann Publishers, Amsterdam.

Fry, H. (2019), *Hello World: How to Be Human in the Age of the Machine*, Transworld Publishers, London, UK.

Gillespie, T. (2014), “The Relevance of Algorithms”, in Gillespie, T., Boczkowski, P.J. and Foot, K.A. (Eds.), *Media Technologies*, The MIT Press, pp. 167–194.

Goodall, D. (2009), “Owned, Bought and Earned Media”, *ALL THAT IS GOOD*, 2 March, available at: https://danielgoodall.wordpress.com/2009/03/02/owned-bought-and-earned-media/ (accessed 12 July 2021).

Graefe, A., Haim, M., Haarmann, B. and Brosius, H.-B. (2018), “Readers’ perception of computer-generated news: Credibility, expertise, and readability”, *Journalism*, Vol. 19 No. 5, pp. 595–610.

Gunaratne, J., Zalmanson, L. and Nov, O. (2018), “The Persuasive Power of Algorithmic and Crowdsourced Advice”, *Journal of Management Information Systems*, Vol. 35 No. 4, pp. 1092–1120.

Haim, M. and Nienierza, A. (2019), “Computational observation: Challenges and opportunities of automated observation within algorithmically curated media environments using a browser plug-in”, *Computational Communication Research*, Vol. 1 No. 1, pp. 79–102.

Helberger, N. (2019), “On the Democratic Role of News Recommenders”, *Digital Journalism*, Vol. 0 No. 0, pp. 1–20.

Jobin, A., Ienca, M. and Vayena, E. (2019), “The global landscape of AI ethics guidelines”, *Nature Machine Intelligence*, Vol. 1 No. 9, pp. 389–399.

Julier, G. (2017), *Economies of Design*, SAGE.

Jung, H., Oh, C., Hwang, G., Oh, C.Y., Lee, J. and Suh, B. (2019), “Tell Me More: Understanding User Interaction of Smart Speaker News Powered by Conversational Search”, *Extended Abstracts of the 2019 CHI Conference on Human Factors in Computing Systems*, Association for Computing Machinery, New York, NY, USA, pp. 1–6.

Kalpokas, I. (2019), *Algorithmic Governance: Politics and Law in the Post-Human Era*, Springer International Publishing, Cham, available at: https://doi.org/10.1007/978-3-030-31922-9 (accessed 14 October 2020).

Kaptein, M., De Ruyter, B., Markopoulos, P. and Aarts, E. (2012), “Adaptive Persuasive Systems: A Study of Tailored Persuasive Text Messages to Reduce Snacking”, *ACM Transactions on Interactive Intelligent Systems*, Vol. 2 No. 2, pp. 1-25.

Kay, M., Matuszek, C. and Munson, S.A. (2015), “Unequal Representation and Gender Stereotypes in Image Search Results for Occupations”, *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, ACM, New York, NY, USA, pp. 3819–3828.

Kim, H.S., Yang, S., Kim, M., Hemenway, B., Ungar, L. and Cappella, J.N. (2019), “An Experimental Study of Recommendation Algorithms for Tailored Health Communication”, *Computational Communication Research*, Vol. 1 No. 1, pp. 103–129.

Kitchin, R. (2014), *The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences*, SAGE Publications, Los Angeles, California.

Kitchin, R. (2017), “Thinking critically about and researching algorithms”, *Information, Communication & Society*, Vol. 20 No. 1, pp. 14–29.

Koene, A., Perez, E., Carter, C.J., Statache, R., Adolphs, S., O'Malley, C., Rodden, T. and McAuley, D. (2015), "Ethics of Personalized Information Filtering", in Tiropanis, T., Vakali, A., Sartori, L. and Burnap, P. (Eds.), *Internet Science*, Springer International Publishing, pp. 123–132.

Krafft, T.D., Gamer, M. and Zweig, K.A. (2019), "What did you see? A study to measure personalization in Google's search engine", *EPJ Data Science*, Vol. 8 No. 1, p. 38.

Liang, T.-P., Robert, L., Sarker, S., Cheung, C.M.K., Matt, C., Trenz, M. and Turel, O. (2021), "Artificial intelligence and robots in individuals' lives: how to align technological possibilities and ethical issues", *Internet Research*, Vol. 31 No. 1, pp. 1–10.

Liu-Thompkins, Y., Maslowska, E., Ren, Y. and Kim, H. (2020), "Creating, Metavoicing, and Propagating: A Road Map for Understanding User Roles in Computational Advertising", *Journal of Advertising*, Vol. 49 No. 4, pp. 394–410.

Loecherbach, F. and Trilling, D. (2020), "3bij3 – Developing a framework for researching recommender systems and their effects", *Computational Communication Research*, Vol. 2 No. 1, pp. 53–79.

Lomborg, S. and Kapsch, P.H. (2020), "Decoding algorithms", *Media, Culture & Society*, Vol. 42 No. 5, pp. 745–761.

Matz, S.C., Kosinski, M., Nave, G. and Stillwell, D.J. (2017), "Psychological targeting as an effective approach to digital mass persuasion", *Proceedings of the National Academy of Sciences*, Vol. 114 No. 48, pp. 12714–12719.

Mayer, R.E. (2014), *The Cambridge Handbook of Multimedia Learning*, Cambridge University Press.

McLuhan, M. (1964), *Understanding Media: The Extensions of Man*, 1st MIT Press ed., MIT Press, Cambridge.

McQuail, D. (2008), “Models of Communication”, *The International Encyclopedia of Communication*, American Cancer Society, available at: https://doi.org/10.1002/9781405186407.wbiecm089 (accessed 15 May 2020).

Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016), “The ethics of algorithms: Mapping the debate”, *Big Data & Society*, Vol. 3 No. 2, pp. 1–21.

Morley, J., Floridi, L., Kinsey, L. and Elhalal, A. (2020), “From What to How: An Initial Review of Publicly Available AI Ethics Tools, Methods and Research to Translate Principles into Practices”, *Science and Engineering Ethics*, Vol. 26 No. 4, pp. 2141–2168.

van Noort, G., Himelboim, I., Martin, J. and Collinger, T. (2020), “Introducing a Model of Automated Brand-Generated Content in an Era of Computational Advertising”, *Journal of Advertising*, Vol. 49 No. 4, pp. 411–427.

Obermeyer, Z., Powers, B., Vogeli, C. and Mullainathan, S. (2019), “Dissecting racial bias in an algorithm used to manage the health of populations”, *Science*, Vol. 366 No. 6464, pp. 447–453.

Ohme, J. (2021), “Algorithmic social media use and its relationship to attitude reinforcement and issue-specific political participation – The case of the 2015 European immigration movements”, *Journal of Information Technology & Politics*, Vol. 18 No. 1, pp. 36–54.

Pariser, E. (2011), *The Filter Bubble: How the New Personalized Web Is Changing What We Read and How We Think*, Penguin.

Perloff, R.M. (2017), *The Dynamics of Persuasion: Communication and Attitudes in the 21st Century*, Sixth edition., Routledge, Taylor & Francis Group, New York.

Proffitt, B. (2011), *The PayPal Official Insider Guide to Selling with Social Media: Make Money through Viral Marketing*, Pearson Education.

Rader, E. and Gray, R. (2015), “Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed", *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, presented at the 33rd Annual ACM Conference, ACM Press, Seoul, Republic of Korea, pp. 173–182.

Reaves, S. (1995), "The Unintended Effects of New Technology: (And Why we Can Expect More)", *Visual Communication Quarterly*, Vol. 2 No. 3, pp. 11–24.

Ricci, F. (2015), *Recommender Systems Handbook*, Springer Science + Business Media, New York, NY.

Richards, J.I. and Curran, C.M. (2002), "Oracles on 'Advertising': Searching for a Definition", *Journal of Advertising*, Vol. 31 No. 2, pp. 63–77.

Rigotti, E. and Rocci, A. (2006), "Towards a definition of communication context.", *Studies in Communication Sciences*, Vol. 6 No. 2, pp. 155–180.

Rodgers, S. and Thorson, E. (Eds.). (2012), *Advertising Theory*, Routledge, New York.

Rose, J. and MacGregor, O. (2021), "The Architecture of Algorithm-driven Persuasion", *Journal of Information Architecture*, Vol. 6 No. 1, pp. 7–40.

Savage, C.W. (2019), "Managing the Ambient Trust Commons: The Economics of Online Consumer Information Privacy", *Stanford Technology Law Review*, Vol. 22, pp. 95–162.

Schiff, D., Rakova, B., Ayesh, A., Fanti, A. and Lennon, M. (2020), "Principles to Practices for Responsible AI: Closing the Gap", *ArXiv:2006.04707 [Cs]*, available at: http://arxiv.org/abs/2006.04707 (accessed 12 July 2021).

Schneider, M.J., Jagpal, S., Gupta, S., Li, S. and Yu, Y. (2017), "Protecting customer privacy when marketing with second-party data", *International Journal of Research in Marketing*, Vol. 34 No. 3, pp. 593–603.

Seeber, I., Waizenegger, L., Seidel, S., Morana, S., Benbasat, I. and Lowry, P.B. (2020), "Collaborating with technology-based autonomous agents: Issues and research opportunities", *Internet Research*, Vol. 30 No. 1, pp. 1–18.

Seeger, A.-M. and Heinzl, A. (2018), "Human Versus Machine: Contingency Factors of Anthropomorphism as a Trust-Inducing Design Strategy for Conversational Agents", in Davis, F.D., Riedl, R., vom Brocke, J., Léger, P.-M. and Randolph, A.B. (Eds.), *Information Systems and Neuroscience*, Springer International Publishing, Cham, pp. 129–139.

Sinclair, J. (2020), "Magazines and Advertising in the Digital Age", *The Handbook of Magazine Studies*, John Wiley & Sons, Ltd, pp. 105–119.

Sonderman, J. and Tran, M. (2013), "The definition of 'sponsored content'", *American Press Institute*, 14 November, available at: https://www.americanpressinstitute.org/publications/reports/white-papers/the-definition-of-sponsored-content/ (accessed 23 February 2020).

Stephen, A.T. and Galak, J. (2012), "The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace", *Journal of Marketing Research*, Vol. 49 No. 5, pp. 624–639.

Strycharz, J., Noort, G. van, Smit, E. and Helberger, N. (2019), "Protective behavior against personalized ads: Motivation to turn personalization off", *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, Vol. 13 No. 2.

Sundar, S.S. (2020), "Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII)", *Journal of Computer-Mediated Communication*, Vol. 25 No. 1, pp. 74–88.

Sundar, S.S., Jia, H., Waddell, T.F. and Huang, Y. (2015), "Toward a Theory of Interactive Media Effects (TIME)", in Sundar, S.S. (Ed.), *The Handbook of the Psychology of Communication Technology*, John Wiley & Sons, Ltd, pp. 47–86.

Sunstein, C.R. (2009), *Republic.Com 2.0*, Princeton University Press.

Susser, D. (2019), “Invisible Influence: Artificial Intelligence and the Ethics of Adaptive Choice Architectures”, *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, Association for Computing Machinery, New York, NY, USA, pp. 403–408.

Sutanto, J., Palme, E., Tan, C.-H. and Phang, C.W. (2013), “Addressing the Personalization-Privacy Paradox: An Empirical Assessment from a Field Experiment on Smartphone Users”, *MIS Quarterly*, Vol. 37 No. 4, pp. 1141–1164.

Tafesse, W. (2020), “YouTube marketing: how marketers’ video optimization practices influence video views”, *Internet Research*, Vol. 30 No. 6, pp. 1689–1707.

Taylor, D.G., Davis, D.F. and Jillapalli, R. (2009), “Privacy concern and online personalization: The moderating effects of information control and compensation”, *Electronic Commerce Research*, Vol. 9 No. 3, pp. 203–223.

Thompson, N.C., Ge, S. and Sherry, Y.M. (2021), “Building the algorithm commons: Who discovered the algorithms that underpin computing in the modern enterprise?”, *Global Strategy Journal*, Vol. 11 No. 1, pp. 17–33.

Thorson, E. and Rodgers, S. (2019), “Advertising Theory in the Digital Age”, *Advertising Theory*, 2nd ed., Routledge.

Tufekci, Z. (2014), “Engineering the public: Big data, surveillance and computational politics”, *First Monday*, Vol. 19 No. 7, available at: https://doi.org/10.5210/fm.v19i7.4901 (accessed 12 May 2020).

Tufekci, Z. (2015), “Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency Symposium Essays”, *Colorado Technology Law Journal*, No. 2, pp. 203–218.

Turel, O., Matt, C., Trenz, M., Cheung, C.M.K., D’Arcy, J., Qahri-Saremi, H. and Tarafdar, M. (2019), “Panel report: the dark side of the digitization of the individual”, *Internet*

Research, Vol. 29 No. 2, pp. 274–288.

Vakratsas, D. and Ambler, T. (1999), "How Advertising Works: What Do We Really Know?", *Journal of Marketing*, Vol. 63 No. 1, pp. 26–43.

Van den Broeck, E., Zarouali, B. and Poels, K. (2019), "Chatbot Advertising Effectiveness: when does the message get through?", *Computers in Human Behavior*, Vol. 98, pp. 150–157.

Vesanen, J. (2007), "What is personalization? A conceptual framework", *European Journal of Marketing*, Vol. 41 No. 5/6, pp. 409–418.

Vesanen J. and Raulas M. (2006), "Building bridges for personalization: A process model for marketing", *Journal of Interactive Marketing*, Vol. 20 No. 1, pp. 5–20.

Vasudevan, K. (2020), "Design of Communication: Two Contexts for Understanding How Design Shapes Digital Media", *Journalism & Mass Communication Quarterly*, Vol. 97 No. 2, pp. 453–468.

Voorveld, H.A.M. and Araujo, T. (2020), "How Social Cues in Virtual Assistants Influence Concerns and Persuasion: The Role of Voice and a Human Name", *Cyberpsychology, Behavior, and Social Networking*, Vol. 23 No. 10, pp. 689–696.

Voorveld, H.A.M., van Noort, G., Muntinga, D.G. and Bronner, F. (2018), "Engagement with Social Media and Social Media Advertising: The Differentiating Role of Platform Type", *Journal of Advertising*, Vol. 47 No. 1, pp. 38–54.

Willson, M. (2017), "Algorithms (and the) everyday", *Information, Communication & Society*, Vol. 20 No. 1, pp. 137–150.

Yeung, K. (2017), "'Hypernudge': Big Data as a mode of regulation by design", *Information, Communication & Society*, Vol. 20 No. 1, pp. 118–136.

Yun, J.T., Segijn, C.M., Pearson, S., Malthouse, E.C., Konstan, J.A. and Shankar, V. (2020), "Challenges and Future Directions of Computational Advertising Measurement

Systems”, *Journal of Advertising*, Vol. 49 No. 4, pp. 446–458.

Zarouali, B., Boerman, S. and de Vreese, C. (2021), “Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale)”, *Telematics and Informatics*, Vol. 62, p. 101607.

Zarouali, B., Dobber, T., De Pauw, G. and de Vreese, C. (2020a), “Using a Personality-Profiling Algorithm to Investigate Political Microtargeting: Assessing the Persuasion Effects of Personality-Tailored Ads on Social Media”, *Communication Research*.

Zarouali, B., Poels, K., Ponnet, K. and Walrave, M. (2020b), “The influence of a descriptive norm label on adolescents’ persuasion knowledge and privacy-protective behavior on social networking sites”, *Communication Monographs*, Vol. 88 No. 1, pp. 5–25.

Zhong, B. (2021), *Social Media Communication: Trends and Theories*, John Wiley & Sons.

Zuiderveen Borgesius, F.J., Möller, J., Kruikemeier, S., Fathaigh, R.Ó., Irion, K., Dobber, T., Bodo, B. and de Vreese, C. (2018), “Online Political Microtargeting: Promises and Threats for Democracy”, *Utrecht Law Review*, Vol. 14 No. 1, pp. 82–96.

Figures

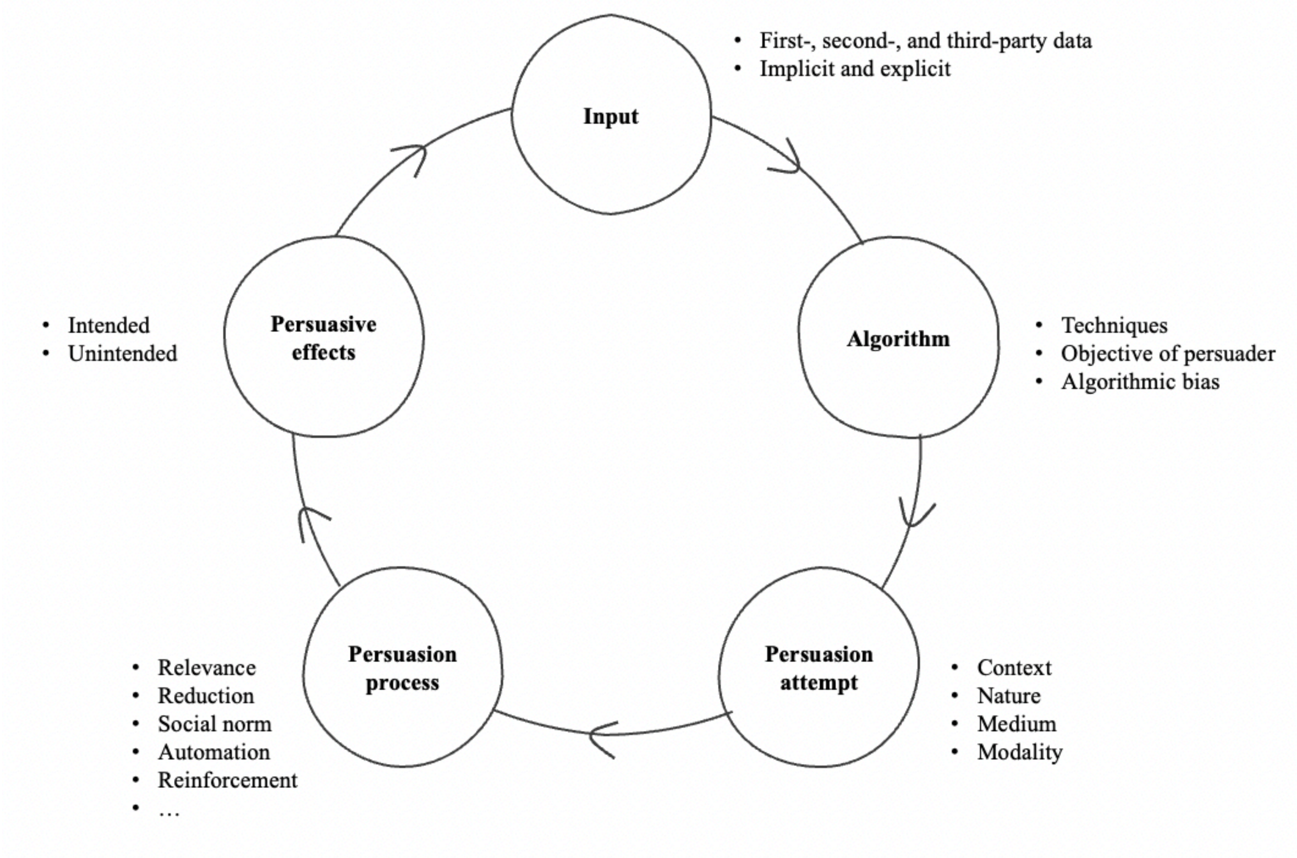

Figure 1. The Algorithmic Persuasion Framework

Tables

Table I. Future research agenda based on the APF

Framework component	Research questions
Input	How are the type of input data (*implicit data or explicit data*) affecting how people are persuaded by algorithmic-mediated communication? How are the type of input data (*first-, second-, or third-party data*) affecting how people are persuaded by algorithmic-mediated communication?
Algorithm	How are different algorithmic techniques (*prioritization, association, classification, and filtering*) contributing to different forms of algorithmic persuasion? How are the *objectives of the persuader* translated into algorithmic code(s), and how does this result in how people are impacted by algorithmic-mediated communication?
Persuasion attempt	How does the *context* (e.g., political communication, health communication, advertising, news, etc.) affect how/the extent to which people are persuaded by algorithmic-mediated content? How does the *nature* (organic vs. paid content) affect how/the extent to which people are persuaded by algorithmic-mediated content? How does the *medium* (e.g., mobile, laptop, smart speaker, etc.) affect how/the extent to which people are persuaded by algorithmic-mediated content? How does the *modality* (e.g., audiovisual content, auditory content, etc.) affect how/the extent to which people are persuaded by algorithmic-mediated content?
Underlying process	Can the underlying mechanisms (i.e., *relevance, reduction, social norm, automation, and reinforcement*) explain how algorithmic persuasion occurs? Which additional mechanisms are relevant in explaining algorithmic persuasion?
Persuasion effects	What are the *intended effects* of algorithmic-mediated communication, and to what extent and under which conditions are they occurring? What are the *unintended effects* of algorithmic-mediated communication, and to what extent and under which conditions are they occurring?
Full framework	How to *set up studies* in which all the components from the APF are included in the research design? How to investigate the *circular dynamic* of the APF? How are the *feedback loops* in the APF affecting algorithmic persuasion? How are all the different components of the APF *interrelated*?