Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale)

Zarouali, B.; Boerman, S.C.; de Vreese, C.H.

Link to publication

Citation for published version (APA):
Zarouali, B., Boerman, S. C., & de Vreese, C. H. (2021). Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale). *Telematics and Informatics*, *62*, [101607]. https://doi.org/10.1016/j.tele.2021.101607
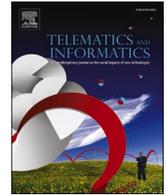
Contents lists available at ScienceDirect

Telematics and Informatics

journal homepage: www.elsevier.com/locate/tele

ELSEVIER

Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale)

Brahim Zarouali [*], Sophie C. Boerman, Claes H. de Vreese

Amsterdam School of Communication Research, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands

A B S T R A C T

Online media platforms are increasingly using algorithms to select and present relevant information to their audiences. This highlights the importance of exploring whether people are aware of algorithmic content recommendations. Although some studies have already investigated *algorithmic awareness*, no standardized instrument has been developed yet to assess this construct. In this study, we therefore developed and validated the Algorithmic Media Content Awareness Scale (AMCA-scale). This scale contains four underlying dimensions: 1) users' awareness of content filtering, 2) users' awareness of automated decision-making, 3) users' awareness of human-algorithm interplay, and 4) users' awareness of ethical considerations. In validating the scale, results revealed strong psychometrics properties. The AMCA-scale was also successfully tested for three different online platforms: Facebook, YouTube, and Netflix, showing its robustness over different environments. Based on these findings, we conclude that the AMCA-scale offers scholars a valid, reliable and robust tool to measure algorithmic awareness.

1. Introduction

Algorithms are increasingly being utilized in a large number of industries, including our contemporary data-driven media landscape (Lee, 2018). Online platforms tailor the selection and presentation of online content (e.g., video content on YouTube, news feed content on social media platforms, audio content on Spotify, etc.) to the individual user using *recommender systems*, i.e., algorithms that decide which content to present to whom based on criteria such as past behavior (Ricci et al., 2015). Users are thus receiving partially distinct streams of online content, initially based on their own (and/or those of similar others) behavioral choices, and subsequently, further personalized by online platforms' algorithmic determinations of what should be prioritized and for what purpose (e.g., to increase ad revenues, time spent on platform, user satisfaction, etc.).

Recently, an increased scholarly attention has been given to the perceptions of and responses towards algorithmic-curated content in online media. In this stream of research, two opposite perceptions have been found. On the one hand, studies present evidence pointing toward the tendency of *algorithmic aversion*, which refers to people's preference to rely on human decisions rather than algorithmic ones (e.g., Dietvorst et all., 2015; Yeomans et al., 2019). On the other, support has been found for *algorithmic appreciation,* where people prefer algorithmic to human judgments (e.g., Logg et al., 2019; Thurman et al., 2019). Although some studies found that these perceptions depend on other factors such as context and individual characteristics (e.g., Araujo et al., 2020; Lee, 2018), the

* Corresponding author.
E-mail address: b.zarouali@uva.nl (B. Zarouali).

research in this area is still very sparse, unconnected, and inconclusive (Jussupow et al., 2020; Schwienbacher, 2020). These inconsistences may be due to differences in the conceptual and operational definitions of algorithmic perceptions. As a response to this, this study takes an important and much-needed step to introduce and develop a measurable construct that lies at the basis of algorithmic perceptions: *algorithmic awareness*. As theoretically argued by Dinev and Hu (2007), having awareness of a new technology is a crucial first step (and precondition) to motivate people to become more cognizant about the technology and form perceptions and opinions about it. Therefore, algorithmic awareness might of significant theoretical value in deriving meaningful and appropriate conclusions about algorithmic perceptions.

Aside from the theoretical relevance of algorithmic awareness, the concept is also of societal importance. On the one hand, being aware of algorithmic recommendations on online platforms might encourage online users to make more critical reflections and decisions regarding the content they are being presented on these platforms (Eslami et al., 2015). On the other, a lack of algorithmic awareness might contribute to major societal problems, such as the spread of mis- and disinformation, the proliferation of filter bubbles, an increased susceptibility to data-driven manipulation, and the reinforcement of stereotypes, inequalities and discrimination (e.g., Eubanks, 2017; Mohamed et al., 2020; Pariser, 2011; Susser, 2019). In the light of this reasoning, we emphasize the importance and urgent need of investigating algorithmic awareness among online media users in a valid and robust way.

However, a validated self-report instrument for algorithmic awareness (or knowledge) has not been developed yet; most existing studies simply relied on their own operationalizations, hereby self-creating questions that fit their own research context (e.g., Bucher, 2017; Cotter & Reisdorf, 2020; Eslami et al., 2015; Gran et al., 2020; Min, 2019; Rader & Gray, 2015). Therefore, guided by the theoretical and practical justifications addressed above, we developed and validated the 13-item *Algorithmic Media Content Awareness scale (AMCA-scale)*. This standardized scale allows scholars to gain insights into people's ability to make proper sense of algorithms in online platforms. During the development of the scale, we applied it to three different platforms (i.e., Facebook, YouTube, and Netflix) to ascertain its usefulness for different online (social) environments. As a reliable and robust instrument, the AMCA-scale allows researchers to compare people's algorithmic awareness between studies and between contexts with respect to their similarities and differences. By this means, we contribute to the accurate and reliable measurement of users' algorithmic perceptions, which may lead to more valuable contributions to scientific theory, public policy and societal debate about the interplay between algorithms and users in a social platformed society.

2. Review of the literature

2.1. Defining algorithmic media content awareness

Algorithms are automated instructions to transform input data into a desired output (Gillespie, 2014). They are important technological ingredients in the architecture of online platforms and are being used to automatically filter enormous amounts of information to offer personalized content, services, and advertisements to the end users (van Dijck, Poell, & de Waal, 2018). In this study, we define algorithmic media content awareness (AMCA) as *the extent to which people hold accurate perceptions of what algorithms do in a particular media environment, as well as their impact on how users consume and experience media content.* Some examples of mediated environments that use algorithms are: filtered newsfeed posts on social media platforms, filtered product offerings on e-commerce websites, filtered video overview on streaming platforms, filtered search engine results, etc. Although these examples have different dynamics from a technical point-of-view, they do share the common tactic of algorithmic content curation, i.e., the use of algorithms (in some way) to select and presents relevant subsets of a large corpus of content to users (Rader & Gray, 2015).

An important first step in scale development consists of reviewing theory and research to pre-specify the structure and meaning of the construct (Carpenter, 2018). This means finding all the subdimensions of the construct (i.e., the breadth of the construct), and creating conceptual definitions for them. Based on a thorough review of the literature (see next section), we propose five dimensions for the algorithmic awareness scale: 1) awareness of content filtering, 2) awareness of automated decision-making, 3) awareness of human-algorithm interplay, and 4) awareness of algorithmic persuasion, and 5) awareness of ethical considerations. All the conceptual definitions of these subdimensions are provided in Table 1. In the next section, we elaborate on these dimensions.

2.2. Dimensions of algorithmic awareness

The first dimension of algorithmic awareness relates to being aware that algorithms are being used to *filter* specific content to people that are most likely to be interested. The explosive growth and variety of content available on the Web introduced a pressing need to reduce a very large set of content into a smaller consideration sets to avoid choice overload and to maximize user relevance

Table 1
Overview of the five initially proposed dimensions of AMCA and their conceptual definitions.

Dimensions	Conceptual definition
Content filtering	Being aware that algorithms are used to *tailor* media content to specific people based on online data
Automated decision-making	Being aware that algorithms are used to make *automated decisions* in tailoring media content to people
Human-algorithm interplay	Being aware that one's *individual behavior* influences which content algorithms decide to present
Algorithmic persuasion	Being aware that algorithms are used to *influence* users' attitudes and behaviors
Ethical considerations	Being aware of the *ethical concerns* of algorithmic-recommended content

(Jannach et al., 2010; Ricci et al., 2015; Schreiner et al., 2019). This has encouraged online platforms to increasingly rely on algorithmic filtering. In doing so, platforms determine the interests, desires, and needs of each user through a substantial amount of datafied user signals, and based on this, personalize or filter content recommendations to align with the interests of the recipient (van Dijck et al., 2018). Although algorithms are widely used, they remain largely covert and opaque to the end users (Beer, 2017; Gillespie, 2014; Lee, 2018). Indeed, studies have shown that people are often characterized by a lack of awareness of algorithmic content curation (e.g., Cotter & Reisdorf, 2020; Eslami et al., 2015; Powers, 2017). This filtering awareness plays an important role in changing users' orientations toward and behaviors on the online platform (Bucher, 2017). Based on this line of reasoning, we argue that algorithmic awareness starts with a fundamental conception that algorithms are used to filter online content.

The second important dimension in algorithmic awareness is *automated-decision making.* Traditionally, humans played an important role in selecting which content to provide to the end user, driven by their professionals and institutional norms and values (DeVito, 2017; Jhaver et al., 2018; Lee, 2018; Logg et al., 2019; Thurman et al., 2019). A key example of this is the news selection and presentation on online news websites, based on the editors' professional and institutional journalistic norms and selection criteria. So, deciding which media content to present was most of the time based on human-centered, knowledge-based assessments (Newell & Marabelli, 2015). These days, platforms increasingly replace these human-based selections with algorithm-driven selection (Diakopoulos, 2019; Jussupow et al., 2020; van Dijck et al., 2018). Thus, platform algorithms are designed to automate human judgments in a highly efficient and optimized way. Being aware of this automated decision making process is an important step to understand how algorithms shape online environments.

People's awareness of the *interplay* between algorithms and users is the third dimension. This means that a platform's selection of which content to present is not only the result of a certain algorithmic logic, but is also shaped by users interacting with these coded environments (Beer, 2017; Vasudevan, 2020; Willson, 2017). As algorithms work in tandem with user input (Bucher, 2018), they should not be seen as "isolated entities"; rather, people continuously shape and re-articulate these algorithms by the choices they make online (Gillespie, 2014; Diakopoulos, 2019; Kitchin, 2017). Thus, by their own behaviors and interactions, people are making algorithms meaningful and able to succeed in doing what they were designed to do. It is important that people are aware of the fact that one's own behavior influences which media content the algorithms decide to show to them. This particular awareness would indicate that people know that algorithms are not only what developers make of them, but also what we make out of them (Fry, 2019; Gillespie, 2014).

As a fourth dimension, we present *algorithmic persuasion.* More than ever, online platforms are competing for the attention of their users in a highly competitive and changing media environment (Andrejevic, 2013). In this battle for attention, platforms use algorithms to induce various forms of persuasion: attitude change, clicking behavior, user engagement, time spent on platform, etc. (Ricci et al., 2015; Siles et al., 2019; Zarouali et al., 2020). The latter depends a great deal on the business model of the platform, or put differently, the ways in which economic value gets created and captured (Gomez-Uribe and Hunt, 2016; van Dijck et al., 2018). But the general tendency here is that algorithms are increasingly creating online persuasive architectures that can influence the online choices of media users in directions preferred by the choice architect (Susser, 2019; Yeung, 2017). For instance, Netflix algorithmically structures its platform to influence the habits and choices of end users regarding how they explore, seek, find and buy content (Siles, Espinoza-Rojas, Naranjo, & Tristán, 2019). Therefore, it is important that end users are not only aware of media content being filtered (see first dimension), but additionally, also realize that this filtering is usually aimed at positively influencing their attention, engagement, and decision making.

Fifth and finally, algorithmic awareness consists of a more sophisticated dimension, which can be referred to as *ethical considerations*. Based on the literature review, we found three important ethical concerns in relation to algorithmic-mediated content: i) privacy intrusion, ii) lack of transparency, iii) algorithmic bias (see Bozdag, 2013; Koene et al., 2015; Saurwein, Just, & Latzer, 2015; Tufekci, 2015; van Dijck et al., 2018). First, privacy risks are mainly caused by the algorithms' need to collect and store personal data about their users to personalize media content (Friedman, Knijnenburg, Vanhecke, Martens, & Berkovsky, 2015). Privacy could be threatened because algorithms might use private or sensitive information (e.g., political orientation) or information that a user has never disclosed (through data modelling, algorithms can make predictions about people's personal traits; Tufecki, 2015; Bozdag, 2013). Second, when it comes to transparency, the concern is that platforms use algorithms to filter content based on inputs and outputs that are neither transparent nor obvious to the human observer (Tufekci, 2015). In addition, the fact that content is being filtered in the first place is also rarely made transparent to users. This lack of transparency makes it difficult for users to understand how their data are being gathered and used by platforms (Koene et al, 2015), hinder them in identifying potential manipulations (Susser et al., 2019), and could prompt them to blindly (dis)trust algorithmic decision-making (van Drunen, Helberger, & Bastian, 2019). Third, algorithmic bias can be considered as a major issue in algorithmic recommended content (Bozdag, 2013; Bucher, 2018; Willson, 2017). As argued by Bozdag (2013), humans usually affect the design of the algorithms, and sometimes also manually influence the filtering process after the algorithm has been designed, which installs human biases. In addition, biases or existing ideologies, prejudices, and inequalities can also slip into the datasets that are used to train the algorithms (Hargittai, 2020; van Dijck et al., 2018). Therefore, it is important for people to be aware that algorithmic decisions are neither "neutral" nor "objective", but a reflection of societal biases.

2.3. Aim of the study

The purpose of this study is to develop an instrument to measure the construct of Algorithmic Media Content Awareness (AMCA), based on the *a priori* set of five relevant dimensions that were identified based on a close inspection of the literature. We aim to develop and validate the AMCA-scale in three distinct phases, in which we (1) systematically identify a pool of items relating to the five

subdimensions of algorithmic media content awareness; (2) develop the AMCA-scale to assess the face and content validity of these items; and (3) validate the scale through data derived from a national representative sample. In this last phase, we aim to divide the total sample in three subsamples representing three different (social) media platforms: Facebook, YouTube, and Netflix. Each sub-sample served a different purpose (i.e., an exploratory sample, a confirmatory sample, and a replication sample). Specific details about
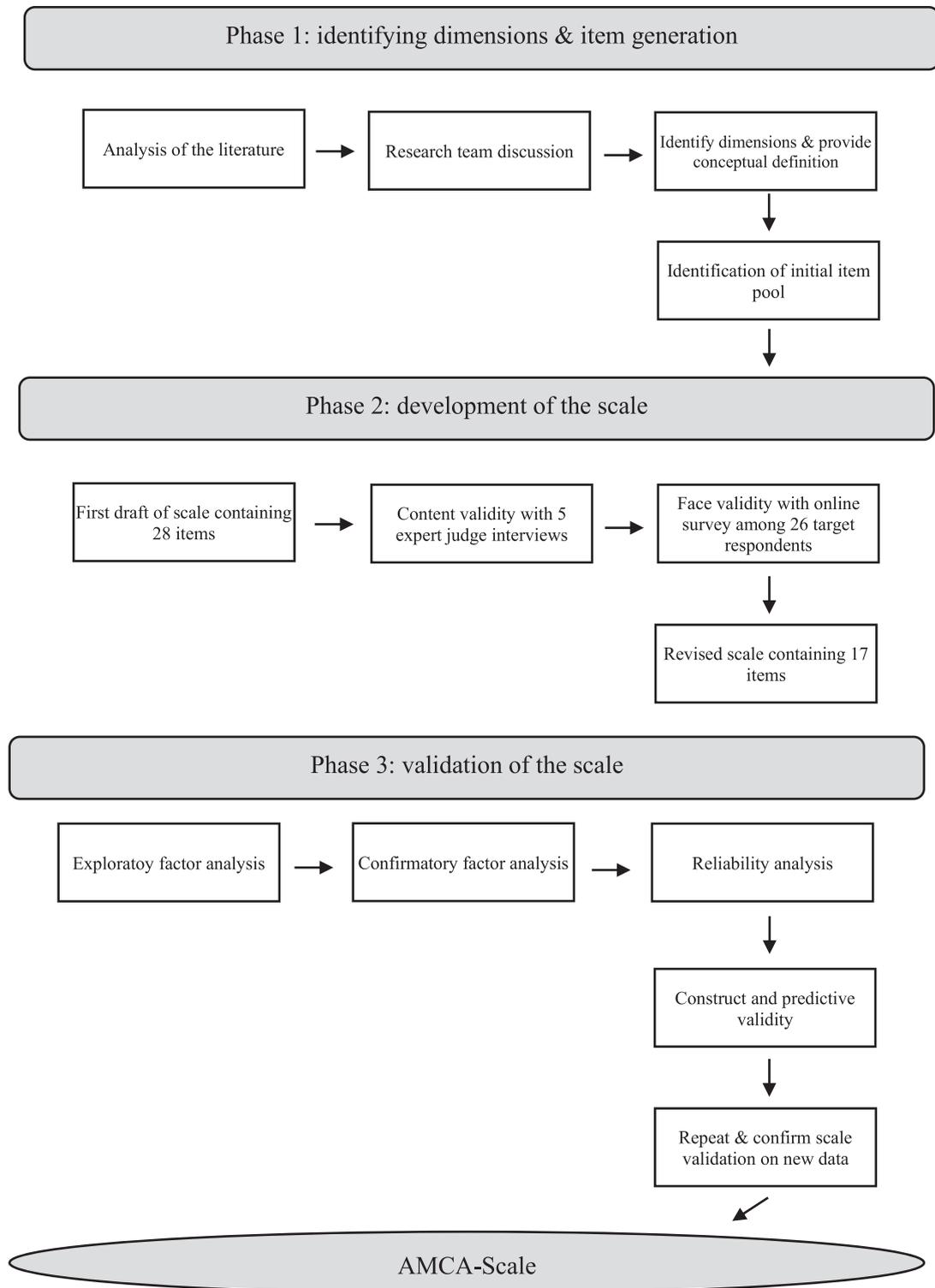

Fig. 1. An overview of all the phases and steps in the development and validation of the AMCA-scale.

this analytical strategy can be found in the section below ("Phase 3"). In Fig. 1, a visual overview is presented of all the steps involved.

3. Methodology

3.1. Phase 1: Identifying dimensions & item generation

After having identified the five subdimensions in the literature, which we provided them with an (initial) formal conceptual definition (Table 1), we proceeded with articulating potential items based on a deductive method. This means that items were generated for the five dimensions based on a careful literature review and a close inspection of existing scale and indicators (Boateng et al., 2018). In this step, we followed item wording guidelines in determining the exact item phrasings, in order to obtain concise, clear, distinct items that reflect the chosen conceptual definitions (Carpenter, 2018; DeVellis, 2016). This initial list of items was carefully refined to eliminate wordings that were still somewhat complex or ambiguous. This iterative process of item generation finally resulted in a pool of 28 items measured on a 5-point rating scale ranging from *1 ("not at all aware")* to *5 ("completely aware")*. The statement that is being used to introduce these items is: *"Please indicate to which extent you are aware of the following statements about algorithms in media content"*. All items were written and developed in English. In Phase 3 (i.e., validating the items among respondents in The Netherlands), the items were translated to Dutch by means of a professional translation service.

3.2. Phase 2: Development of the scale

3.2.1. Content validity

Once the items were generated, we aimed to test for content validity, which is usually done through the evaluation of expert judges (Boateng et al., 2018; Polit & Beck, 2006). We interviewed five academic experts, consisting of two communication scholars (subject-matter researchers), a sociologist (i.e., researcher on the role of algorithms in news), a computer scientist (research on recommender systems and AI), and a research methodologist (expert in survey design). After explaining the construct and dimensions, the experts were asked to provide feedback about their interpretation of the construct, their evaluation of the underlying dimensions, followed by their judgements of each individual item, as well as the instrument as a whole (Carpenter, 2018). In this phase, they were asked to verbalize any comment (e.g., improvement or omission of an item, rewording to make an item more clear, etc.). To quantify this content validity step, we used the Content Validity Index on the item level (I-CVI; see Lynn, 1986; Polit & Beck, 2006). The experts were asked to rate each item on a 4-point ordinal scale (*1 = not relevant; 2 = item is in need of such revision that it would no longer be relevant; 3 = relevant item but needs minor alteration; 4 = very relevant*). For each item, the I-CVI is computed as the number of experts giving either a score of 3 or 4, divided by the total number of experts (Polit & Beck, 2006). Following the criteria of Lynn (1986), we omitted the items with an I-CVI-score lower than 1, which led to the elimination of 11 items. Based on the experts' comments, we rewrote some of the retained items. This process resulted in a final pool of 17 items (algorithmic filtering: 4 items; automated decision-making: 3 items; human-algorithm interplay: 4 items; algorithmic persuasion: 3 items; ethical considerations: 3 items).

3.2.2. Face validity

A crucial step in scale development is ensuring face validity, which implies that the prospective respondents should understand all the items, as well as evaluate the items as appropriate to the targeted construct and objectives (Haynes et al., 1995). To assess face validity, we conducted a pre-test among the target respondents of the instrument (Boateng et al., 2018; Carpenter, 2018). For this purpose, we conducted an online survey among 16 respondents, hereby presenting open-ended questions allowing them to modify, clarify and improve each item. More precisely, they were asked to complete the entire scale, and while doing so, provide suggestions and feedback about whether the specific items and response categories were clear, understandable, and grammatically correct, had no spelling mistakes, and did not involve jargon. Finally, respondents were asked whether they had suggestions to improve the overall scale to fit the objectives of the study. Although the respondents indicated some small mistakes, they all agreed on the dimensions, as well as the clarity and understandability of all the items. In sum, this feedback on the items revealed good face validity from the perspective of the target audience.

3.3. Phase 3: Validation of the scale

3.3.1. Sample

Respondents were randomly recruited among members of a research panel (consisting of approx. 80.000 – 100.000 members) from a renowned market research company. In total, 6000 respondents were randomly invited. The final sample consisted of 2106 respondents, which reveals a response rate of 35%. This random sample was representative for the Dutch population in terms of age, gender, educational level and region. The data collection took place in July and August 2019, and all respondents received an incentive (bonus points) for completing the questionnaire. The participants had a mean age of 54 (*SD* = 15.59 years), and 52% of them were male. Around 31% had a lower education level (no education or primary education), 50% medium levels of education (secondary education), and 19% higher education (bachelor, master or doctoral degree).

3.3.2. Analytical strategy

To validate the AMCA-scale, we followed the recommended steps for scale development (Boateng et al., 2018; Carpenter, 2018; Worthington & Whittaker, 2006). Exploratory factor analysis (EFA) was used as a means of determining the number and nature of the

underlying factors in the scale, followed by a confirmatory factor analysis (CFA) to confirm whether the data fit this extracted factor structure. Then, we conducted a reliability analysis, and tested the construct validity. This was followed by a replication, which includes a confirmatory factor analysis, assessment of internal consistency reliability and construct validation (Hinkin, 1998). The final step of the scale development process, which took place *four months later*, was assessing the test–retest reliability. Psychometric analyses were carried out in the statistical program R.

The total sample consisted of three subsamples: a Facebook subsample ($n = 688$), a Netflix subsample ($n = 741$), and a YouTube subsample ($n = 677$). Participants randomly received the scale items related to one of these platforms. All items were formulated the same across the three conditions, with only the platform (Facebook vs. YouTube vs. Netflix) and content type varying (personalized posts vs. personalized videos vs. personalized movies). We made use of subsamples for two main reasons. First, it allowed to check whether the scale is applicable to different platforms. Second, it enabled us to adopt a three-step-approach: the first subsample served to provide guidance about item selection and factor reduction (exploratory sample), the second subsample to confirm this data structure and assess reliability and construct validity (confirmatory sample), and the third subsample to repeat the previous procedures (Noar, 2003; DeVellis, 2016). We selected the Netflix subsample to serve as the exploratory sample, the Facebook subsample as the confirmatory sample, and the YouTube subsample as the replication sample. This selection happened at random.

3.3.3. Exploratory factor analysis

Prior to conducting the EFA, we visually inspected the correlation matrix between all variables (Pearson) in the exploratory subsample, i.e., the Netflix sample. All correlations were above the threshold 0.30, and none of them were too high (<0.90), which indicates a reduced likelihood of multicollinearity issues (Field, Miles, & Field, 2012). We performed the Bartlett's test of sphericity to investigate the factorability of the data and a Kaiser-Meyer-Olkin (KMO) test to measure the sampling adequacy. The KMO test showed sampling adequacy (KMO = 0.99), and Bartlett's test confirmed that the data is appropriate for EFA ($\chi^2 = 16461.24\ df = 120, p < .001$). Then, we proceeded with a factor analysis with principal axis factoring, and used an oblique rotation strategy (i.e., Promax). Oblique rotation was chosen because it allows underlying factors to be correlated, which is expected in multidimensional scales (Bowman & Goodboy, 2020). Parallel analysis was used to determine the number of factors to retain. The EFA resulted in a four-factor solution (and not five, as theorized *a priori*), which explained 74% of the variance. All items had high communalities (Table 2), except for one specific item (item HAI3), which was considerably lower than all the other items (communality of 0.49). We decided to drop this item as it did not load clearly on one specific factor as well. The factor structure and their respective factor loadings can be found in Table 2 (highest loadings are in bold).

Looking at the factor loading clusters, we can conclude that the factor solution only extracted four factors that were *a priori* theoretically defined. The first four items (FIL1-FIL4) have their main loading on one factor, which represents the 'algorithmic filtering' dimension. Then, items ADM1-ADM3 clearly load on a separate factor as well, which therefore refers to the automated decision-making dimension. Then, the loadings of items HAI1, HAI2 and HAI4 are also loading high on one factor, which is 'human-algorithm interplay' (HAI3 has been removed, because of reasoning addressed above). When it comes to items ALP1- ALP3, a different conclusion can be drawn. These items are *not* clearly loading onto one factor. Based on this exploratory solution, we decided to omit these three items from further analyses. Finally, we have the three final items ETC1-ETC3 that form a cluster around a separate factor

Table 2
Exploratory factor analysis of AMCA.

Item	Factor 1	Factor 2	Factor 3	Factor 4	h^2
Filtering					
Item FIL1	0.02	**0.75**	0.11	0.07	0.86
Item FIL2	0.19	**0.66**	0.05	0.01	0.76
Item FIL3	0.03	**0.78**	0.08	0.08	0.89
Item FIL4	0.09	**0.40**	0.17	0.28	0.76
Automated decision-making					
Item ADM1	0.08	0.01	0.02	**0.78**	0.75
Item ADM2	0.06	0.13	0.08	**0.67**	0.82
Item ADM3	0.01	0.29	0.17	**0.51**	0.86
Human-algorithm interplay					
Item HAI1	0.03	0.26	**0.61**	0.06	0.84
Item HAI2	0.09	0.12	**0.71**	−0.05	0.78
Item HAI3	**0.36**	0.04	0.14	0.21	0.49
Item HAI4	0.02	0.17	**0.76**	0.09	0.81
Algorithmic persuasion					
Item ALP1	0.19	0.09	**0.49**	0.18	0.79
Item ALP2	0.19	**0.27**	0.32	0.20	0.83
Item ALP3	0.25	**0.36**	0.28	0.08	0.84
Ethical considerations					
Item ETC1	**0.73**	0.12	−0.10	0.22	0.71
Item ETC2	**0.75**	0.06	0.14	−0.04	0.78
Item ETC3	**0.47**	−0.01	0.37	0.10	0.77

Note. Based on the Netflix subsample (i.e., the exploratory sample). Factor loadings lower than 0.50 are not presented (except for FI2, where the highest loading is under this threshold).

as well, which is the dimension 'ethical considerations'.

An issue worth mentioning is the presence of some cross-loadings among (i.e., FIL4, ADM3, and ETC3 – see Table 2). Although cross loadings are common in social science scale development, they do deserve careful attention (Howard, 2016). One approach would be to delete these three items, but however, one should be cautious when using cross-loadings for item deletion as to not compromise scale length and factor structure (Worthington & Whittaker, 2006). In this case, item deletion would lead to two factors ending up with two items (i.e., automated decision making, and ethical considerations). It has been argued that factors should ideally contain no less than 3 items to represent a "clean" factor structure (Osborne et al., 2008). So rather than omitting these three items, we decided to retain them as part of their intended factor.

3.3.4. Confirmation and assessment of reliability and validity

CFA was used to assess how well this four-factor structure fits the observed data. This was done by using data of the confirmatory sample, i.e., the Facebook subsample. As a model configuration, we tested a hierarchical model. This model tests the idea that a second-order factor can account for relations between the four AMCA factors, and thus would suggest that all factors are related to this higher order factor (Noar, 2003). Therefore, a good model fit of a hierarchical model indicates that the four factors can be used as separate scales, as well as together as an entire overarching meaningful scale.

We used maximum likelihood estimation to conduct the CFA. To evaluate the fit, we used the conventional fit indices , i.e., the chi-squared divided by the degrees of freedom ($\chi2/df$), the comparative fit index (CFI), the Tucker–Lewis index (TLI), the standardized root mean squared residual (SRMR) and the root mean squared error of approximation (RMSEA). The fit indices revealed a good fit to
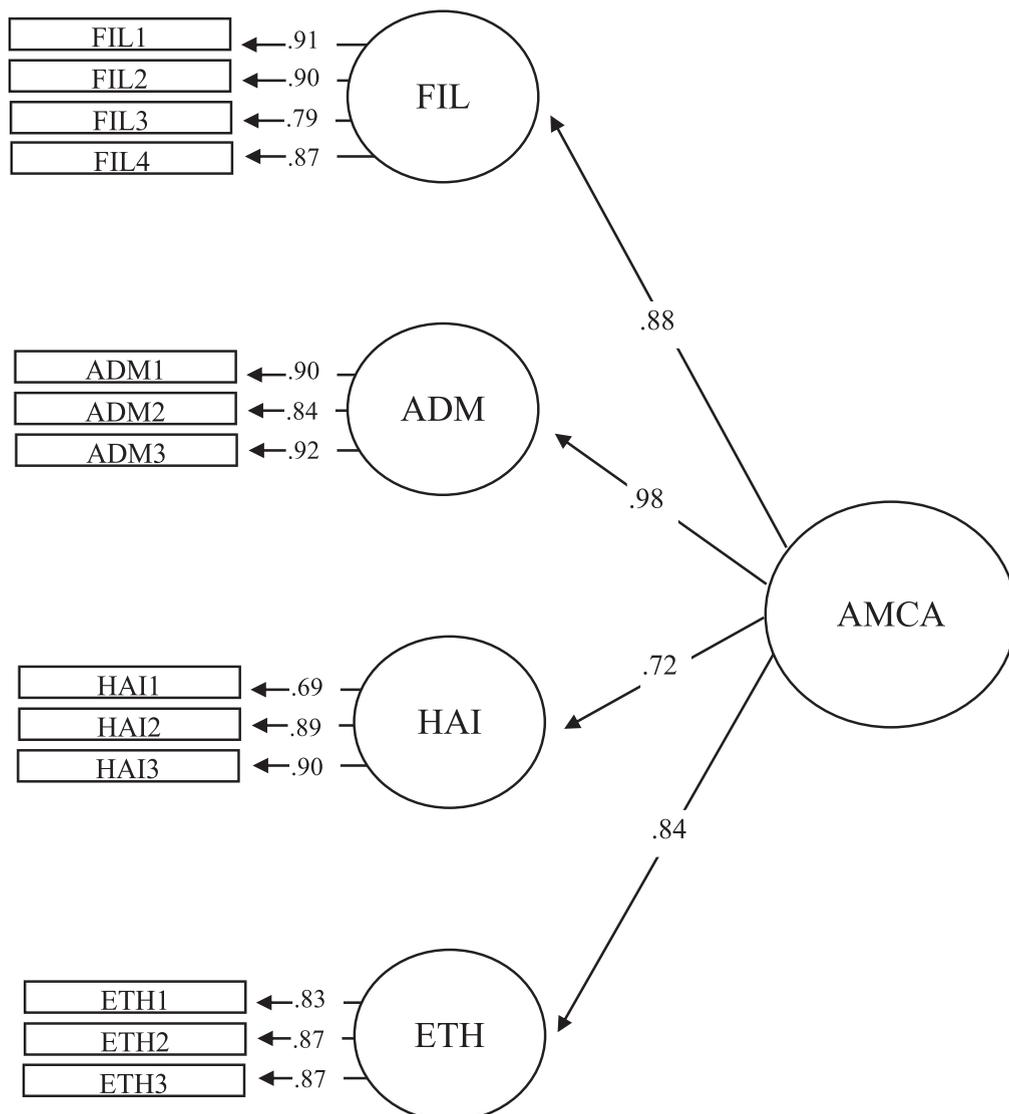
Fig. 2. Higher-order measurement model of the AMCA-scale (Facebook subsample).

the data (χ^2/df = 3.46, RMSEA = 0.060, SRMR = 0.02, CFI = 0.99, TLI = 0.98), without any re-specification or modification of the model. The model reveals high standardized factor loadings, ranging from 0.69 to 0.92 (all significant; p < 0.001) (see Fig. 2). Altogether, we retain this well-fitting second-order model, showing that the four factors are subscales can be tested individually as well as summed together into one larger scale (Noar, 2003).

Next, we assessed the internal consistency of the scale for the confirmatory subsample (see Table 4). The following Cronbach alpha values were found: for the filtering dimension, the internal reliability was α = 0.92, for the automated decision-making dimension α = 0.92, for human-algorithm interplay α = 0.86, and for individual implications α = 0.89. These numbers indicate high reliability of the AMCA-subdimensions.

Then, we assessed construct validity of the confirmatory subsample, referring to the degree to which a scale measures what it is intended to measure. It consists of *convergent* and *discriminant* validity. Convergent validity is evaluated by the size of the standardized factor loadings, the average variance explained (AVE) and the construct reliability (Bagozzi & Heatherton, 1994; Bagozzi & Yi, 1988). AVE is calculated as the mean variance extracted by the item loading on a specific construct, and as a rule, a value higher than 0.50 is considered to be an indicator of adequate convergence. Given that all standardized loadings in the CFA model were strong and significant (see earlier), the AVE for all factors are well beyond the threshold of 0.50 (all AVE were between 0.70 and 0.79; Fornell & Larcker, 1981), and the reliability of all constructs is greater than 0.70, we conclude that convergent validity is established. A rigorous way to assess discriminant validity is to compare the AVE values for any two dimensions with the square root of the correlation estimate between these two dimensions. If the AVE is higher than the squared correlation estimate, then discriminant validity is demonstrated (Fornell & Larcker, 1981; Hair et al., 2014). The subsample passed this test for all dimensions. Therefore, this provides evidence that the subdimensions reflect independent constructs, rather than a single construct.

Finally, we also tested predictive validity, which refers to the extent to which a score on the AMCA-scale can predict scores on a criterion measure. As argued by Dinev and Hu (2007), awareness is an important first step that leads to more knowledge about a technology. Therefore, we expect AMCA to predict algorithmic knowledge. To test this, we measured respondents' knowledge about algorithms by presenting them 7 true/false statements about common algorithmic misperceptions. The responses were then dummy coded to determine if a correct answer was given (correct vs. incorrect). Algorithmic knowledge was then computed by adding up all the correct answers (scores ranging from 0 to 7). A higher score indicated greater algorithmic knowledge (i.e., less misconceptions). An OLS regression revealed that AMCA significantly predicted people's algorithmic knowledge (b = 0.84, p < .001; F(1,685) = 190.17, p < .001; R^2 = 0.22).

Table 3
The AMCA-scale with all the final items.

Dimension	Items	Response categories
	Please indicate to which extent you are aware of the following statements about algorithms in media content.	
Content filtering	FIL1: Algorithms are used to recommend [media content] to me on [platform name]	1 = *not at all aware* 5 = *completely aware*
	FIL2: Algorithms are used to prioritize certain [media content] above others	
	FIL3: Algorithms are used to tailor certain [media content] to me on [platform name]	
	FIL4: Algorithms are used to show someone else see different [media content] than I get to see on [platform name]	
Automated decision-making	ADM1: Algorithms are used to show me [media content] on [platform name] based on automated decisions	1 = *not at all aware* 5 = *completely aware*
	ADM2: Algorithms do not require human judgments in deciding which [media content] to show me on [platform name]	
	ADM3: Algorithms make automated decisions on what [media content] I get to see on [platform name]	
Human-algorithm interplay	HAI1: The [media content] that algorithms recommend to me on [platform name] depend on my online behavior on that platform	1 = *not at all aware* 5 = *completely aware*
	HAI2: The [media content] that algorithms recommend to me on [platform name] depend on my online behavioral data	
	HAI3: The [media content] that algorithms recommend to me on [platform name] depend on the data that I make available online	
Ethical considerations	ETH1: It is not always transparent why algorithms decide to show me certain [media content] on [platform name]	1 = *not at all aware* 5 = *completely aware*
	ETH2: The [media content] that algorithms recommend to me on [platform name] can be subjected to human biases such as prejudices and stereotypes	
	ETH3: Algorithms use my personal data to recommend certain [media content] on [platform name], and this has consequences for my online privacy	

Note. Depending on the platform, the [platform name] included Netflix, Facebook or YouTube, and [media content] said movies, posts, or videos.

3.3.5. Replication

As argued by Hinkin (1998), the replication part should include a confirmatory factor analysis, assessment of internal reliability and construct validation. For this, the replication sample was used, i.e., the YouTube subsample. We used CFA to test the hierarchical model tested earlier by using maximum likelihood estimation. With this subsample, result show even better goodness-of-fit-indices (χ^2/df = 3.46, RMSEA = 0.048, SRMR = 0.01, CFI = 0.99, TLI = 0.99). All factor loadings are very high with values between 0.82 and 0.92 (all are significant; $p < .001$). Therefore, these results show a very good fit of the second-order model in the YouTube condition.

In testing the reliability and construct validity, we found similar findings as in the Facebook subsample. The filtering dimension had a Cronbach's α of 0.92, the automated decision-making dimension α = 0.94, for human-algorithm interplay α = 0.92, and ethical considerations α = 0.88. For construct validity, we again assessed convergent and discriminant validity. For the reason that all standardized factor loadings in the CFA model were high (between 0.82 and 0.92) and significant, the AVE estimates all exceeded 0.50 (the AVE for the dimensions were all between 0.70 and 0.81), and reliability estimates all exceeded 0.70 (see earlier), it can be concluded that convergent validity is established. As for the discriminant validity, we used the same approach in the previous section. The YouTube subsample also passed the test since all AVE values for any two dimensions were higher than the square root of the correlation between the two dimensions, and thus establishing discriminant validity. Finally, predictive validity was tested. OLS regressions indicated that the AMCA-scale could predict respondents' algorithmic knowledge ($b = 0.83, p < .001; F(1,675) = 183.43, p < .001; R^2 = 0.21$).

Having successfully replicated the main psychometric properties of the AMCA-scale, we present the descriptive estimates of the finalized scale in Table 4. All the finalized items of the AMCA can be found in Table 3.

3.3.6. Test-retest reliability

Finally, we evaluated the temporal stability of the AMCA-scale, i.e., test–retest reliability. It reflects the variation in measurements (at different time intervals) taken by an instrument on the same subject under the same conditions (Koo and Li, 2016). Researchers vary in how they quantify test–retest reliability: some use the intra-class correlation (ICC), while others use the Pearson product-moment correlation (Boateng et al., 2018). We used ICCs because, unlike Pearson correlations, ICCs do not only assess covariation but also agreement between the measurements at the two time intervals (Weir, 2005). ICC values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability (Koo and Li, 2016).

The re-administration of the survey containing the AMCA-scale took place four months after the first assessment. In total, 1288 respondents filled out the survey in this second wave. Two-way random effects ICCs revealed: ICC (content filtering) = 0.76, with [95% CI 0.724–0.783]; ICC (automated decision-making) = 0.75, with [95% CI 0.720–0.780]; ICC (human-algorithm interaction) = 0.72 with [95% CI 0.684–0.747]; ICC (ethical considerations) = 0.70 with [95% CI 0.661–0.729]. Hence, based on these results, we conclude that there is evidence for the temporal stability of the AMCA-scale.

4. General discussion

As algorithms are increasingly used in the online environment, it is important to be able to measure whether people are aware of this development. In this study, we developed the Algorithmic Media Content Awareness Scale (AMCA-scale) (see Table 3). This scale measures people's awareness of the usage and consequences of algorithms for the media content on different online platforms (e.g., Netflix, YouTube, Facebook).

The AMCA-scale was developed in three phases (see Fig. 1 for an overview). Based on the first two phases, we anticipated to develop five dimensions in phase three, but the EFA revealed that the items load on four dimensions (four-factor structure). Thus, the final scale encompasses four components: 1) users' awareness of content filtering, 2) users' awareness of automated decision-making, 3) users' awareness of human-algorithm interplay, and 4) users' awareness of ethical considerations.

During the development of the scales, we ensured its content and face validity, and during the validation process, the scales had strong psychometrics properties such as a good reliability, convergent and discriminant validity, and test–retest reliability. In addition, CFA showed that the four factors can be combined to measure one overall construct (i.e., hierarchical model test). This means that researchers can use the scales separately, to measure people's awareness of one or more of the specific dimensions (as they see fit in the context of their research). In addition, the four factors can be combined as an overall summated scale of people's awareness of algorithmic media content. This approach offers scholars flexibility in using the AMCA-scale either as a general construct, or as a multidimensional instrument with separate measures.

The AMCA-scale was successfully tested for three different media platforms: Netflix, Facebook, and YouTube. Based on high similarities in estimates of the scales between the platforms, we expect that the AMCA-scale can also (validly and reliably) be used for other platforms. However, as each platform has its own characteristics and features, further research is needed to test how the AMCA-scale performs in other algorithmic-mediated contexts, such as news platforms, e-commerce platforms, music streaming platforms, etc. When using the scale in these different contexts, two small adaptations need to be done: changing the "[platform name]" in the items, as well as the type of "[media content]" (see Table 3).

This study offers an important methodological contribution to the literature of algorithmic awareness. As already addressed in the introduction, no validated measure of algorithmic awareness has been developed yet. Instead, studies simply relied on self-created items fitting their specific research context (e.g., Bucher, 2017; Cotter & Reisdorf, 2020; Eslami et al., 2015; Gran et al., 2020; Min, 2019). When we compare the AMCA-scale results to these existing studies, we come to a similar conclusion: respondents have -in

general- a rather low awareness of algorithms (see Table 4). So the AMCA-scale results are mostly in line with previous algorithmic awareness studies. However, compared to these previous studies, the AMCA-scale has some important advantages: i) it presents a valid measure based on a rigorous scale development process; ii) it taps into different subdimensions of algorithmic awareness; and iii) it was shown to be a robust scale across different platforms. These advantages should lead to more methodological standardization and consistency among studies, and thus, allow scholars to compare the findings of studies with respect to their differences and similarities in a more systematic way. In the end, this should lead to more valuable contributions to public policy and debate about the role of algorithms in our society.

4.1. Theoretical implications

The AMCA-scale, as presented in this study, has important theoretical implications. As already discussed in the introduction, the literature on algorithmic perceptions in online platforms is still very sparse and inconclusive. This made it difficult to make theory-based generalizations to a wider population, nor compare the results across different studies. The AMCA-scale now offers a valid and reliable measure to assess people's algorithmic awareness, which is a key component in the formation of attitudes and perceptions toward this new technology (Dinev and Hu, 2017). That is, the concept of awareness is central in understanding human responses and behaviors in many domains, including new technologies. In the next paragraphs, we discuss four theoretical implications of the AMCA-scale.

First, the AMCA measure can lead to theoretical advancements in the area of algorithmic perceptions and attitudes. That is, AMCA can be used as key predictor or antecedent when it is expected to cause or predict a specific outcome. For instance, algorithmic awareness might predict people's trust perceptions toward online algorithms in online platforms. Alternatively, algorithmic awareness can also be seen as a theoretical concept in its own right, where future research could find out how variations in awareness can be explained (causal or correlational) by other variables (e.g., tech savviness, media use, etc.). In addition, in the context of algorithmic perceptions, AMCA can also advance the literature of algorithmic appreciation vs. algorithmic aversion (see introduction). So far, many studies in this area have revealed findings that are not entirely conclusive (Jussupow et al., 2020; Schwienbacher, 2020). We argue that AMCA can be positioned as an important theoretical construct in forming perceptions of appreciation or aversion. As argued by Dinev and Hu (2017), having awareness of a new technology is a crucial first step (and precondition) to make sure people can form perceptions and opinions about it. Therefore, the AMCA-scale might contribute to the literature by offering the valid operationalization of an important theoretical construct that lies at the heart of the formation of algorithmic perceptions.

Second, awareness has been addressed as an important moderator in determining people's perceptions and responses towards new technologies (Abubakar, 2013). Therefore, we argue that AMCA-scale could function as an individual-level moderator, for instance in experimental research. This means that algorithmic awareness could serve as an important variable that influences the magnitude of algorithmic media effects or alters algorithmic perceptions and attitudes. So, from a theoretical perspective, this could contribute to our understanding that effects and perceptions in the area of algorithmic media content vary between people with different levels of algorithmic awareness.

Third, awareness about a technology has also been addressed as an important theoretical part of various digital literacies (e.g., media literacy, data literacy, etc.) (Hosman and Pérez Comisso, 2020). Therefore, future research might consider to investigate how algorithmic awareness relates to digital literacy and competences, and focus on how important digital literacy is in establishing algorithmic awareness (e.g., do differences in digital literacy translate into differences in algorithmic awareness?). In addition, the AMCA-scale can also be used to examine the overall "algorithmic literacy" of a given population, how this literacy changes over time (longitudinal designs) or differs across cultures or countries, and what personal or situational characteristics can predict or influence this literacy. Importantly, it still remains to be tested whether the AMCA-scale can be applied in different cultures and countries in the exact same way (and if not, future research could look into the specific adjustments that need to be made). But all in all, the scale can be an important instrument to identify (vulnerable) people with low literacy levels that may be in need of user empowerment, media education, and maybe even regulatory protection. Altogether, the AMCA-scale can deliver insights that contribute to the understanding of the broader field of digital literacy.

4.2. Limitations

Finally, the study has some limitations as well, that could be addressed in future research. First, although the four dimensions that are included in the AMCA-scale are based on an extensive literature review and previous empirical work, we cannot guarantee that this

Table 4
Summary of descriptive indices (M, SD, and Cronbach's α).

	Netflix subsample (N = 741)			Facebook subsample (N = 688)			Youtube subsample (N = 677)			Total sample N = 2106		
	M	*SD*	*α*	*M*	*SD*	*α*	*M*	*SD*	*α*	*M*	*SD*	*α*
FIL	2.94	1.43	0.90	3.18	1.30	0.92	3.14	1.34	0.94	3.09	1.36	0.92
ADM	2.87	1.42	0.93	3.17	1.33	0.92	3.07	1.32	0.92	3.03	1.36	0.92
HAI	2.90	1.42	0.93	3.09	1.24	0.86	3.18	1.32	0.92	3.05	1.34	0.90
ETH	2.71	1.30	0.88	3.16	1.26	0.89	3.02	1.24	0.88	2.96	1.28	0.89

Note: FIL = Content filtering, ADM = automated decision-making, USI = human-algorithm interplay, ETH = ethical considerations.

is an exhaustive list of dimensions. There may be other dimensions that are part of the awareness construct that are not included in the scale. Future conceptual studies may uncover relevant components that can be added to the scale. Second, we did not have clear empirical evidence to underpin one of the five *a priori* hypothesized dimensions, i.e., algorithmic persuasion. This could indicate that this theoretical subscale is not distinctive enough (i.e., invariant to one or more of the other dimensions), or that we did not have the appropriate items to capture this construct correctly. Since we believe that this dimension can be of significant relevance for future research, we encourage scholars to examine how this concept should be operationalized and measured in a valid and reliable way. Third, this study was largely centered around AMCA's five-dimensional assumption, which was based on a literature review. In future research, it might be worthwhile to use qualitative measures such as in-depth interviews, focus groups, or think-aloud sessions to explore possible missing dimensions. Such an exploratory process would help to gain insights into the true potential breadth of the concept AMCA.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Abubakar, F.M., 2013. The moderating effect of technology awareness on the relationship between UTAUT constructs and behavioural intention to use technology. Aust. J. Business Manage. Res. 3 (2), 14–23.

Andrejevic, M. (2013). Infoglut: How too much information is changing the way we think and know. Routledge.

Araujo, T., Helberger, N., Kruikemeier, S., de Vreese, C.H., 2020. In AI we trust? Perceptions about automated decision-making by artificial intelligence. AI SOC. 35 (3), 611–623. https://doi.org/10.1007/s00146-019-00931-w.

Bagozzi, R.P., Heatherton, T.F., 1994. A general approach to representing multifaceted personality constructs: application to state self-esteem. Structural Equation Modeling: A Multidisciplinary Journal 1 (1), 35–67. https://doi.org/10.1080/10705519409539961.

Bagozzi, R.P., Yi, Y., 1988. On the evaluation of structural equation models. JAMS 16 (1), 74–94. https://doi.org/10.1007/BF02723327.

Beer, D., 2017. The social power of algorithms. Inform. Commun. Soc. 20 (1), 1–13. https://doi.org/10.1080/1369118X.2016.1216147.

Boateng, G.O., Neilands, T.B., Frongillo, E.A., Melgar-Quiñonez, H.R., Young, S.L., 2018. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front. Public Health 6, 149.

Bowman, N.D., Goodboy, A.K., 2020. Evolving considerations and empirical approaches to construct validity in communication science. Ann. Int. Commun. Assoc. 44 (3), 219–234. https://doi.org/10.1080/23808985.2020.1792791.

Bozdag, E., 2013. Bias in algorithmic filtering and personalization. Ethics Inf. Technol. 15 (3), 209–227. https://doi.org/10.1007/s10676-013-9321-6.

Bucher, T., 2017. The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms. Inform. Commun. Soc. 20 (1), 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Bucher, T., 2018. If...then: Algorithmic power and politics. Oxford University Press, New York.

Carpenter, S., 2018. Ten steps in scale development and reporting: a guide for researchers. Commun. Meth. Measures 12 (1), 25–44. https://doi.org/10.1080/19312458.2017.1396583.

Cotter, K., Reisdorf, B.C., 2020. Algorithmic knowledge gaps: a new dimension of (digital) inequality. Int. J. Commun. 14, 21.

DeVellis, R.F., 2016. Scale development: Theory and applications, 4th ed. SAGE, Thousand Oaks, Calif.

DeVito, M.A., 2017. From Editors to Algorithms: A values-based approach to understanding story selection in the Facebook news feed. Digital J. 5 (6), 753–773. https://doi.org/10.1080/21670811.2016.1178592.

Diakopoulos, N., 2019. Automating the news: how algorithms are rewriting the media. Harvard University Press, Cambridge, Massachusetts.

Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them err. Journal of Experimental Psychology: General, 144(1), 114–126.

Dinev, T., & Hu, Q. (2007). The Centrality of Awareness in the Formation of User Behavioral Intention toward Protective Information Technologies. Journal of the Association for Information Systems, 8(7).

Eslami, M., Rickman, A., Vaccaro, K., Aleyasen, A., Vuong, A., Karahalios, K., … Sandvig, C. (2015). "I always assumed that I wasn't really that close to [her]": Reasoning about Invisible Algorithms in News Feeds. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15, 153–162.

Eubanks, V., 2017. Automating inequality: How high-tech tools profile, police, and punish the poor, first ed. St. Martin's Press, New York.

Field, A.P., Miles, J., Field, Z., 2012. Discovering statistics using R. Sage, London; Thousand Oaks, Calif.

Fornell, C., Larcker, D.F., 1981. Structural equation models with unobservable variables and measurement error: algebra and statistics. J. Market. Res. 18 (3), 382–388. https://doi.org/10.1177/002224378101800313.

Friedman, A., Knijnenburg, B.P., Vanhecke, K., Martens, L., Berkovsky, S., 2015. In: Recommender Systems Handbook. Springer US, Boston, MA, pp. 649–688. https://doi.org/10.1007/978-1-4899-7637-6_19.

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 T. Gillespie, P. J. Boczkowski, & K. A. Foot (Eds.), Media Technologies (pp. 167–194).

Gomez-Uribe, C.A., Hunt, N., 2016. The Netflix Recommender System: Algorithms, Business Value, and Innovation. ACM Trans. Manage. Inf. Syst. 6, 1–19. https://doi.org/10.1145/2843948.

Gran, A.-B., Booth, P., Bucher, T., 2020. To be or not to be algorithm aware: a question of a new digital divide? Inform. Commun. Soc. 1–18.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., 2014. Multivariate Data Analysis, 7 edition. Pearson, Upper Saddle River, NJ.

Hargittai, E., 2020. Potential Biases in Big Data: Omitted Voices on Social Media. Soc. Sci. Comput. Rev. 38, 10–24. https://doi.org/10.1177/0894439318788322.

Haynes, S. N., Richard, D. C. S., & Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. Psychological Assessment, 7(3), 238–247.

Hinkin, T.R., 1998. A brief tutorial on the development of measures for use in survey questionnaires. Organiz. Res. Meth. 1 (1), 104–121. https://doi.org/10.1177/109442819800106.

Hosman, L., Pérez Comisso, M.A., 2020. How do we understand "meaningful use" of the internet? Of divides, skills and socio-technical awareness. J. Inform. Commun. Ethics Soc. 18, 461–479. https://doi.org/10.1108/JICES-05-2020-0055.

Howard, M.C., 2016. A review of exploratory factor analysis decisions and overview of current practices: what we are doing and how can we improve? Int. J. Human-Computer Interact. 32 (1), 51–62. https://doi.org/10.1080/10447318.2015.1087664.

Jhaver, S., Karpfen, Y., Antin, J., 2018. Algorithmic Anxiety and Coping Strategies of Airbnb Hosts. In: Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, pp. 1–12.

Jussupow, E., Benbasat, I., Heinzl, A., 2020. Why are we averse towards algorithms? A comprehensive Literature Review on Algorithm Aversion. ECIS 2020 Proceedings, 1–18.

Kitchin, R., 2017. Thinking critically about and researching algorithms. Inform. Commun. Soc. 20, 14–29. https://doi.org/10.1080/1369118X.2016.1154087.

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

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J. Chiropr. Med. 15, 155–163. https://doi.org/10.1016/j.jcm.2016.02.012.

Lee, M.K., 2018. Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic management. Big Data & Society 5 (1). https://doi.org/10.1177/2053951718756684.

Logg, J.M., Minson, J.A., Moore, D.A., 2019. Algorithm appreciation: people prefer algorithmic to human judgment. Organ. Behav. Hum. Decis. Process. 151, 90–103. https://doi.org/10.1016/j.obhdp.2018.12.005.

Lynn, M. R. (1986). Determination and Quantification of Content Validity. Nursing Research, 35(6), 382–386.

Min, S.J., 2019. From algorithmic disengagement to algorithmic activism: Charting social media users' responses to news filtering algorithms. Telematics Inform. 43, 101251. https://doi.org/10.1016/j.tele.2019.101251.

Mohamed, S., Png, M.-T., Isaac, W., 2020. Decolonial AI: Decolonial Theory as Sociotechnical Foresight in Artificial Intelligence. Philos. Technol. 33, 659–684. https://doi.org/10.1007/s13347-020-00405-8.

Newell, S., Marabelli, M., 2015. Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of 'datification'. J. Strateg. Inf. Syst. 24 (1), 3–14. https://doi.org/10.1016/j.jsis.2015.02.001.

Noar, S.M., 2003. The role of structural equation modeling in scale development. Struct. Eq. Model.: A Multidisciplinary Journal 10 (4), 622–647. https://doi.org/10.1207/S15328007SEM1004_8.

Osborne, J. W., Costello, A. B., & Kellow, J. T. (2008). Best Practices in Exploratory Factor Analysis. In J. Osborne, Best Practices in Quantitative Methods (pp. 86–99).

Pariser, E., 2011. The filter bubble: what the Internet is hiding from you. Viking, London.

Polit, D.F., Beck, C.T., 2006. The content validity index: are you sure you know what's being reported? critique and recommendations. Res. Nurs. Health 29 (5), 489–497. https://doi.org/10.1002/nur.20147.

Powers, E., 2017. My News Feed is Filtered?: Awareness of news personalization among college students. Digital Journalism 5 (10), 1315–1335. https://doi.org/10.1080/21670811.2017.1286943.

Rader, E., & 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, 173–182.

Ricci, F., Rokach, L., Shapira, B., 2015. In: Recommender Systems Handbook. Springer US, Boston, MA, pp. 1–34. https://doi.org/10.1007/978-1-4899-7637-6_1.

Saurwein, F., Just, N., & Latzer, M. (2015). Governance of algorithms: Options and limitations. Info, 17(6), 35–49.

Schreiner, T., Rese, A., Baier, D., 2019. Multichannel personalization: Identifying consumer preferences for product recommendations in advertisements across different media channels. J. Retail. Consum. Serv. 48, 87–99. https://doi.org/10.1016/j.jretconser.2019.02.010.

Schwienbacher, J. (2020). Reactions on algorithms: A systematic literature review of algorithm aversion and algorithm appreciation [University of Innsbruck]. http://diglib.uibk.ac.at/ulbtirolhs/5099158.

Siles, I., Espinoza-Rojas, J., Naranjo, A., Tristán, M.F., 2019. The Mutual Domestication of Users and Algorithmic Recommendations on Netflix. Communication, Culture and Critique tcz025.

Susser, D., 2019. Invisible Influence: Artificial Intelligence and the Ethics of Adaptive Choice Architectures. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, AIES '19. Association for Computing Machinery, New York, USA, pp. 403–408.

Susser, D., Roessler, B., & Nissenbaum, H. (2019). Technology, autonomy, and manipulation. Internet Policy Review, 8(2).

Thurman, N., Moeller, J., Helberger, N., Trilling, D., 2019. My Friends, Editors, Algorithms, and I: Examining audience attitudes to news selection. Digital Journalism 7 (4), 447–469. https://doi.org/10.1080/21670811.2018.1493936.

Tufekci, Z., 2015. Algorithmic Harms beyond Facebook and Google: Emergent Challenges of Computational Agency Symposium Essays. Colo. Technol. Law J. 2, 203–218.

van Dijck, J., Poell, T., & de Waal, M. de. (2018). The platform society. New York: Oxford University Press.

van Drunen, M. Z., Helberger, N., & Bastian, M. (2019). Know your algorithm: What media organizations need to explain to their users about news personalization. International Data Privacy Law.

Vasudevan, K., 2020. Design of Communication: Two Contexts for Understanding How Design Shapes Digital Media. J. Mass Commun. Quart. 97, 453–468. https://doi.org/10.1177/1077699020916427.

Weir, J.P., 2005. Quantifying Test-Retest Reliability Using the Intraclass Correlation Coefficient and the SEM. J. Strength Cond. Res. 19, 231. https://doi.org/10.1519/15184.1.

Willson, Michele, 2017. Algorithms (and the) everyday. Inform. Commun. Soc. 20 (1), 137–150. https://doi.org/10.1080/1369118X.2016.1200645.

Worthington, Roger L., Whittaker, Tiffany A., 2006. Scale development research: a content analysis and recommendations for best practices. Counseling Psychol. 34 (6), 806–838. https://doi.org/10.1177/0011000006288127.

Yeomans, Michael, Shah, Anuj, Mullainathan, Sendhil, Kleinberg, Jon, 2019. Making sense of recommendations. J. Behav. Dec. Making 32 (4), 403–414. https://doi.org/10.1002/bdm.2118.

Yeung, Karen, 2017. 'Hypernudge': big Data as a mode of regulation by design. Inform. Commun. Soc. 20 (1), 118–136. https://doi.org/10.1080/1369118X.2016.1186713.

Zarouali, B., Dobber, T., De Pauw, G., de Vreese, C., 2020. Using a personality-profiling algorithm to investigate political microtargeting: Assessing the persuasion effects of personality-tailored ads on social. Communication Research, media.