

Technology, Mind, and Behavior • Volume 3, Issue 1

It's Just a Recipe?— Comparing Expert and Lay User Understanding of Algorithmic Systems

Thao Ngo¹, Nicole Krämer²

¹Department of Computer Science and Applied Cognitive Science, Research Training Group “User-Centred Social Media,” University of Duisburg-Essen,

²Department of Computer Science and Applied Cognitive Science, Social Psychology: Media and Communication, University of Duisburg-Essen

Published on: Nov 08, 2021

DOI: 10.1037/tmb0000045

License: [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License \(CC-BY-NC-ND 4.0\)](https://creativecommons.org/licenses/by-nc-nd/4.0/)

ABSTRACT

Algorithmic systems may appear opaque to users. This can hinder users from making informed decisions about the use of such systems. To combat this, explanations intend to make them more transparent. However, explanations are typically informed by the system properties. We argue that they also need to consider the user understanding in order to be more comprehensible to users. To achieve these *user-informed* explanations, this qualitative work aims to (a) compare how experts and lay users understand algorithmic systems and (b) derive implications for creating user-informed explanations. We conducted an expert focus group ($N = 3$) and semistructured in-depth interviews with experts ($N = 10$) and lay users ($N = 11$), including a drawing task. Reflexive thematic analysis by the first author revealed group-specific and common themes: Experts understood algorithms as a decision-making process and were aware of the context dependency of algorithms. Lay users, in turn, understood algorithms as intelligence and as data structuring. They focused on the tangible and visible elements of algorithmic systems. Both groups also understood algorithms as a sequence of actions. The different understandings might be driven by group-specific experiences and purposes to use algorithmic systems. Based on our results, we argue that user-informed explanations could consider the context dependency of algorithmic systems and highlight their limitations.

Keywords: transparency, explanation, qualitative study, thematic analysis, drawing task

Acknowledgment: This work was supported by the Deutsche Forschungsgemeinschaft (DFG) under Grant No. GRK 2167, Research Training Group “User-Centred Social Media.” We acknowledge support by the Open Access Publication Fund of the University of Duisburg-Essen.

Conflict of interest: The authors declare that there is no conflict of interest.

Data availability statement: Interview transcripts are not publicly available due to them containing information that could compromise research participant privacy and informed consent. Use case descriptions, interview guidelines, and codes that support the findings of this study are openly available on the Open Science Framework (OSF): <https://osf.io/72tgn/>.

Correspondence concerning this article should be addressed to Thao Ngo, Department of Computer Science and Applied Cognitive Science, Research Training Group “User-Centred Social Media,” University of Duisburg-Essen, Forsthausweg 2, 47057 Duisburg, Germany thao.ngo@uni-due.de

Algorithmic systems are increasingly applied in crucial online domains ([Adadi & Berrada, 2018](#); [Craglia et al., 2018](#)). The application of complex algorithmic systems has diverse implications for the individual and society. Among others, their application affects the users’ autonomy as well as privacy and concerns social issues of responsibility and accountability of algorithms, that is, the question of who can be held responsible and accountable for the people, data, action, and consequences ([Craglia et al., 2018](#)). Yet, their inner mechanisms often remain opaque to users: Users might not fully understand which personal data are collected and how the systems process them to arrive at specific outputs.

However, users need to understand why and how their personal data are used and what kind of inferences can be made from them. This, ultimately, concerns the users’ autonomy, that is, the users’ ability to act upon their own informed choice and privacy ([Craglia et al., 2018](#)). [Cotter & Reisdorf, 2020](#) argue that users need to understand the inner working of algorithms and the factors that affect them to be able to “make rational judgments about the information that they encounter” (p. 748).

Therefore, previous research has focused on the increase of algorithmic transparency through explanations ([Diakopoulos, 2015](#); [Tintarev & Masthoff, 2007](#)). Explanations can influence various factors positively, such as user trust, transparency perception, or scrutability ([Kunkel et al., 2019](#); [Tintarev & Masthoff, 2007](#)). The explanations are usually informed by the system properties, that is, the model, input, and output, but not necessarily by the (lack of) user knowledge. As a result, the explanations are typically not tailored to the prior users’ knowledge level and might be too difficult or too obvious for users.

Against this background, we intend to contribute to explanations that are tailored to the understanding of lay users. To this end, this work compares the user understanding of experts (here, users with computer science or related background) and lay users. By contrasting the understanding of these user groups, we can infer the differences between the most detailed and technically correct knowledge on algorithmic curation and the average understanding of lay users. The resulting insights can inform explanations that better address the existing knowledge gaps.

Consequently, this work has two research aims: The first aim is to describe and characterize this knowledge gap. Like [Eiband et al., 2018](#), we argue that this gap highlights the differences and the shared understanding of these groups and can help shape explanations. Based on these insights, the second aim is to infer the elements of a system that should be explained to lay users. We define this type of explanation *user-informed explanations*. Here, we do not argue that lay users should be turned into experts, as this is unrealistic and unfeasible. Instead, lay users need to understand algorithmic systems in a way that allows them to make informed decisions about their use.

We chose algorithmic curation as an example to study the differences of how lay users understand algorithmic systems compared to experts. Algorithmic curation is the automated selection, organization, and presentation of information. In social networks and news sites, algorithmic curation personalizes the online content influencing media exposure with the intention to cater relevant information to users ([Diakopoulos, 2015](#)). Prior research in the domain of security and privacy indicates that experts' understanding is multilayered compared to lay users' understanding which is simpler and service oriented ([Kang et al. \(2015\)](#)). However, a concrete analysis of the differences between experts' and lay users' understanding of algorithmic curation is missing to this date. Our research questions were:

- *Research Question 1:* How do experts understand algorithmic curation systems?
- *Research Question 2:* How do lay users understand algorithmic curation systems?
- *Research Question 3:* What are the differences between expert and lay user understanding of algorithmic curation systems?

To address these questions, we chose a qualitative approach using a focus group and semistructured interviews as well as a drawing task. We analyzed our material using reflexive thematic analysis. It is suitable and flexible to approach our research questions as we were interested in interpreting the material to compare the subjective experts' and lay users' answers ([Braun & Clarke, 2006](#), [Braun & Clarke, 2020](#)). It allowed us to identify patterns, so-called themes, that are shared but also distinct to the user groups. We favored this method instead of reliability or codebook approaches as we aimed at establishing an in-depth exploration and characterization of user understanding rather than quantify our material ([Vaismoradi et al., 2013](#)).

We argue that shared understanding can be used as a template for explanations. The differences in understanding reveal the lay users' unawareness and misconceptions of specific system components. This could advance current research on transparency and explainability by determining how an effective explanation should be designed.

Categories and Goals of Explanations

Explanations can contribute to a better understanding of a particular subject. They inform individuals about some “sense of mechanism” and often entail causal relations ([Keil, 2006](#), p. 228). Furthermore, explanations can help individuals with understanding why a certain event occurred and engage in justification for it or predict certain events ([Keil, 2006](#)). Researchers have proposed numerous categories of explanations for algorithmic systems. For instance, model-centric explanations, also referred to as global explanations, provide general information about algorithmic systems. Subject-centric explanations, also known as local explanations, are based on the input data given to an algorithmic system ([Došilović et al., 2018](#); [Edwards & Veale, 2017](#)). A similar distinction is provided by [Friedrich & Zanker, 2011](#): White-box-explanations (how-explanations) describe how an algorithmic system derives a particular outcome based on a specific input, while black-box explanations (why-explanations) justify specific outputs. These categories were extended by what-explanations, which reveal “the existence of algorithmic decision-making” (p. 2), and objective explanations, which highlight the algorithm as unbiased and improving ([Rader et al., 2018](#)). All these explanations rely on the system properties to inform explanations. In other words, they consider the type of the algorithmic model of a system, the input data, or the output of a system. However, the user understanding, that is, what the user knows about an algorithmic system or its inner working, is neglected in these explanations. Thus, we argue that it is not ensured that users comprehend the explanation correctly and can make sense of it.

Research on recommender systems has discussed the explanation goals focusing on user-centric measures, for example, trustworthiness, user satisfaction, scrutability, and transparency ([Balog & Radlinski, 2020](#); [Tintarev & Masthoff, 2007](#)). These goals correlate to a certain extent ([Balog & Radlinski, 2020](#)). While they are essential, we argue that the effectiveness of an explanation is also determined by whether the user can comprehend the explanation at hand. It is, therefore, important to examine whether it increases the users' existing knowledge about the system. This lack of consideration was previously stressed by [Miller \(2018\)](#). Explanations are not merely

“presentation of associations and causes” (p. 7) but contextual. Thus, he emphasized the importance of the user’s perception of the algorithmic systems.

So far, there are a few types of user-informed explanations, that is, explanations consider the users’ knowledge and use them as a base for creating an explanation. User-informed explanations can be independent of the type of content or form. For instance, [Chang et al. \(2016\)](#) combined crowd sourcing and natural language processing to create crowd-based explanations for movie recommendations. These explanations were written and evaluated by users and were perceived as more useful and trustworthy and increased users’ satisfaction. The study highlights the importance of more user-informed explanations and their potential benefits for trustworthiness or satisfaction.

Another user-informed explanation was developed by [Cai et al. \(2019\)](#) for a drawing application. These visual example-based explanations were derived from users’ real drawings. [Cai et al. \(2019\)](#) investigated normative example-based explanations (showing a norm from a certain drawing) and comparative example-based explanations (showing the most similar drawings to the users’ drawing). They found that the normative example-based explanation affected the understanding of the system positively. However, the measurement of the user understanding lacked depth as participants were asked to which degree they understand the system through 1 Likert-scale item. Thus, we argue that the measurement was not able to capture the user understanding extensively. To explore the user understanding in-depth, it can be elicited through conceptualization, such as mental models ([Norman, 1983](#)) or folk theories ([2018](#); [Eslami et al. \(2016\)](#); [Gelman & Legare, 2011](#)).

User Understanding of Technological and Algorithmic Systems

Previous studies on user understanding of how a technological and algorithmic system works have mostly applied qualitative approaches. In the field of cognitive psychology and human-computer interaction, mental models can be defined as cognitive knowledge representations of technological systems. They encompass the subjective understanding of a technological system and might be incomplete and flawed. They are constructed through system interaction ([Norman, 1983](#)). Researchers have claimed that mental models’ alignment with the respective conceptual model of a system is crucial for its comprehension and application ([Asgharpour et al., 2007](#); [Eiband et al., 2018](#); [Norman, 1983](#)).

Few studies have explicitly compared experts' and lay users' mental models. These studies have shown that technical background and expertise play a role in mental models. For instance, [Hmelo-Silver & Pfeffer, 2004](#) found that aquarium experts relied on structural elements to develop mental models. They discovered that lay users relied more on visible features for their mental model. In the field of security and privacy, [Asgharpour et al., 2007](#) have shown that mental models of security risks by experts and lay users differ in their nature. Thus, risk communication needs to address the lay user's mental model and their respective perceptions of relevant risks ([Asgharpour et al., 2007](#); [Jorgensen et al., 2015](#)).

While expertise is playing a role in the nature of a mental model, it does not directly translate into more secure online behavior but rather a higher awareness of possible threats and risks ([Kang et al. \(2015\)](#)). [Renaud et al. \(2014\)](#) argue that, besides the lack of understanding of the technological system, other factors, such as the lack of understanding of the consequences of risks and untrustworthy information sources, as well as personal experience, also contribute to the behavior.

Another branch of research on user understanding of algorithmic systems encompasses the conceptualization of lay user understanding as folk theories, which can be defined as intuitive informal theories ([2018](#); [Gelman & Legare, 2011](#)). Folk theories entail causal relations and can help users explain, interact with, and predict the world. Moreover, they are imprecise and can embody cognitive biases ([Gelman & Legare, 2011](#)). While mental models and folk theories overlap in their definition, folk theories are a looser conceptualization of guiding beliefs and do not strongly adhere to a mechanistic structure ([2018](#)).

Folk theories have been investigated in the context of algorithmic curation on specific social network sites ([DeVito et al., 2017, 2018](#); [Eslami et al. \(2016\)](#)). For the algorithmic curation on Facebook, [Eslami et al. \(2016\)](#) identified 10 different folk theories that differ in the sense of control users have over their social feed. For example, while some users believed that their feed was driven by the number of interactions ("Personal Engagement Theory"), others believed that the feed favored visual content, such as photos and videos ("Format Theory").

A comparison between expert and lay user understanding of algorithmic curation is—to our best knowledge—missing to this date. In this work, we focus on the exploration, description, and characterization of the group-specific user understanding, that is, the users' conception of how algorithmic curation works and what their inner mechanisms are to determine online content. While previous research has also focused on one

specific platform, we were curious whether there were overarching themes of algorithmic curation systems.

Method

The local ethics committee approved the study of the University of Duisburg-Essen. Participants consented to recordings of the sessions. Interviews were transcribed. All identifiable information was anonymized.

Participants

In total, 24 German participants took part in our study ([Table 1](#)). Our sample size for interviews was based on several considerations regarding the data quality and the nature of our topic (user understanding was captured through verbal and visual expressions, interviews between 60 and 80 min including a drawing task) and the scope of our study ([Morse, 2000](#)). Similar previous studies sampled between 20 and 30 participants. Additionally, the sample size was driven by the limited availability of experts. Regarding the focus group, we deemed a group size of three as optimal as we prioritized that experts would have sufficient time and the possibility to reflect and discuss their understanding.

Seven participants identified as female, and 17 as male. Participants' age ranged from 20 to 73 years ($M = 31.67$, $SD = 11.49$). Thirteen participants were considered experts. Participants qualified as an expert if they had a university degree or vocational training in computer science or similar degrees and currently work in the field of computer science. Furthermore, experts had programming skills. Consequently, participants who did not qualify as an expert were considered as lay users. We recruited experts and lay users separately in two recruitment phases, respectively. All participants were recruited through social media platforms, email, and personal contact.

Procedure

We conducted one expert focus group in December 2019 and in-depth semistructured interviews with experts and lay users in early spring 2020. While we were able to conduct the focus group and the first 10 expert interviews face to face, all other interviews were conducted online through video-conference tools. This was due to the restrictions of the COVID-19 pandemic, which included social distancing and restricted personal contact.

Table 1

| Overview of Participants' Demographics | | | | |
|--|--------|-----|----------------------|--------------------------------------|
| ID | Gender | Age | Education background | Professional background |
| Experts in focus group (N = 3) | | | | |
| 1 | M | 28 | PhD | Researcher in network analysis |
| 2 | M | 33 | Master | Researcher in recommender systems |
| 3 | F | 28 | Master | Researcher in computer linguistics |
| Experts in individual interviews (N = 10) | | | | |
| 4 | F | 36 | PhD | Researcher in computer linguistics |
| 5 | M | 26 | BA | Computer science student |
| 6 | M | 36 | PhD | Assistant professor in Security |
| 7 | M | 26 | Vocational training | IT specialist for system integration |
| 8 | F | 31 | PhD | Software developer |
| 9 | M | 59 | Vocational training | Senior security specialist |
| 10 | M | 25 | Master | Researcher in IT-Security |

| | | | | |
|--|---|----|-------------|--------------------------------|
| 11 | M | 25 | Master | Researcher in machine learning |
| 12 | M | 26 | Master | PhD student in IT-Security |
| 13 | M | 31 | Master | Software developer |
| Lay users in individual interviews (N = 11) | | | | |
| 14 | F | 29 | Master | Corporate communication |
| 15 | M | 22 | High school | Law student |
| 16 | M | 20 | High school | Teaching student |
| 17 | M | 27 | High school | Business psychology student |
| 18 | M | 73 | PhD | Retired |
| 19 | F | 24 | Bachelor | Consumer Science student |
| 20 | M | 30 | Master | History student |
| 21 | F | 31 | Master | Human resources |
| 22 | F | 30 | Law degree | Lawyer |
| 23 | M | 30 | Master | Product manager |
| 24 | M | 34 | Master | Accounting & Finance |

Initial Expert Focus Group

We conducted an initial focus group with three experts to explore the expert understanding of algorithmic systems. While we originally planned to elicit a joint expert understanding with the focus group only, we recognized that further in-depth

interviews with experts from different fields and professionals were necessary. The focus group participants advised us to conduct further interviews as expert understanding can be diverse and context dependent. The focus group's procedure adhered to the same structure and questions as the individual semistructured interviews described below. Experts also performed a drawing task in which we asked them to visualize the inner working of a news or social media algorithmic curation system. We assumed that this task could support participants to express their ideas of algorithms' functioning and would force them to be less vague in their answers. In contrast to the individual interviews, the focus group participants engaged in group discussions instead of individual probing. The focus group lasted 2 hr, and each participant received 20 euros as compensation for participation.

Individual Semistructured Interviews

We subsequently interviewed 10 experts and 11 lay participants. At the beginning of each interview, participants were introduced to the topic of algorithmic curation. Here, the interviewer stated that the research focuses on social feeds and news curation. Then, participants were asked to write down all associations they had regarding these use cases. They had to explain their associations and group them afterward. Participants elaborated on each association. The interviewer asked specifically for the inner working as well as capabilities and limits of the algorithmic curation system. After that, the interviewer introduced the drawing task. Participants were asked to draw how the algorithmic curation system works on a sheet of paper. They were instructed to explain their drawing in-depth and encouraged to openly talk about their ideas (e.g., they were reminded that there are no right or wrong answers). In the end, participants were debriefed. As compensation for participation, participants received 10–12 euros, depending on the length of interviews. Experts' interviews lasted around 1 hr, while lay users' interviews lasted around 1 hr and 20 min. Use case descriptions, the interview guideline, and codes are publicly available on OSF: <https://osf.io/72tgn/>.

Thematic Analysis

We applied a *reflexive* thematic analysis to analyze our material. This analysis is a method for identifying, analyzing, and reporting underlying patterns, so-called themes, and poses one type of thematic analysis among many. It is distinct from other thematic analysis approaches, such as coding and codebook reliability thematic analyses. In the reflexive thematic analysis, the analysis of the material is subject to the interviewer's interpretation ([Braun & Clarke, 2006](#), [Braun & Clarke, 2020](#)). [Braun & Clarke, 2020](#) note that “a research team is not required or even desirable for quality” (p. 6). Data

analysis was performed after all interviews were completed and transcribed. The analysis was carried out in MAXQDA2018 by the first author.

The first author extensively familiarized herself with the material and coded it. The analysis was rather inductive than deductive. Codes were constantly refined and revised throughout the analysis and the interpretation process and were formed into overarching themes. She used descriptive and in-vivo coding, which captured participants' voices at the beginning of the analysis. In later stages of the analysis, she mostly applied pattern coding ([Saldaña, 2013](#)). For instance, final code examples of the common theme algorithms as a sequence of actions included: abstract, context dependence, or “theory vs. practice” with subcodes of “algorithms as maths” and “algorithms as a program.”

Researcher Description

The reflexive thematic analysis considers the researcher as an “analytic resource” ([Braun & Clarke, 2020](#), p. 3) to interpret the material at hand. As such, the researcher's background and position affect the analysis and interpretation. The interviewer and coder of this study is a female PhD student who has a background in psychology and human factors. The interviewer can be considered as a layperson with high technical interest and affinity. Additionally, as some participants were contacted through first- and second-degree personal contact, they were known to the interviewer before the study. The relation to these participants was mostly professional; two exceptions were in closer relation to the interviewer.

Results

Reflexive thematic analysis as conducted by the first author revealed four different themes: One theme specific to expert participants, two themes specific to lay participants, and a common theme shared by both groups. The first author consistently noticed in the material that discussions on algorithmic curation systems often resulted in the use of the general term of “algorithms.” As such, the results were extended from algorithmic curation systems to the general understanding of algorithmic systems.

The lay participants' understanding in our sample was characterized by an emphasis on user data and the output of the data. Concerning the questions of how data are processed and how it arrives at a decision, lay participants expressed uncertainty (e.g., L3: “I don't know exactly how algorithms work.”), but still had some elaborated beliefs about the algorithmic model. This difference was also visible in some drawings, for instance, in the drawing of L3 compared to the drawing of E1 ([Figure 1](#)). L3 left out

the algorithmic model and only illustrated his data input (in this case, a tree). In this regard, it is interesting to note that six lay participants mentioned the necessary technical devices for algorithmic systems (server, PC). Expert participants typically abstractly described algorithms without mentioning any devices.

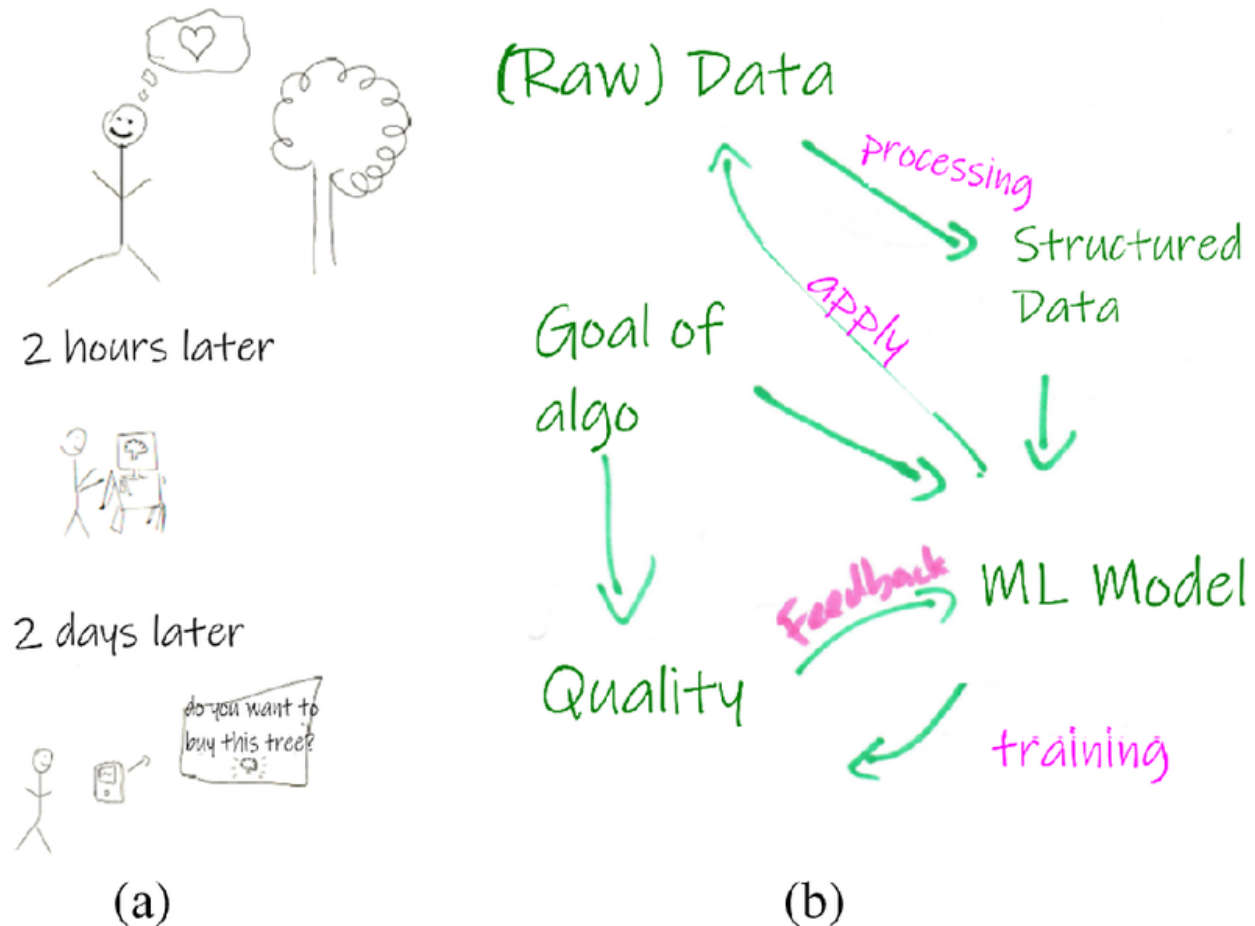


Figure 1
Drawing of P3 and E1

Note. (a) L3 expressed an individual view on algorithms showing the data input (liking a tree”) and the data output (recommendation of a tree 2 days later on his device). (b) E1 had a structural view on algorithms expressing different elements including raw data, structured data, ML model, quality, and goal of algorithms and their relationships among each other.

Experts’ Theme: Algorithms as a Decision-Making Process

The analysis by the first author demonstrated that expert participants viewed *algorithms as a decision-making process*: They described that algorithmic systems are used to solve a predefined problem. The problem was understood as the task an

algorithm has to fulfill. The task, in turn, determines all important elements of the algorithmic system, including the necessary data, the algorithmic model, and the output of the algorithms. Therefore, algorithms are constructed and understood as a tool that “just does statistics” (E8). As such, they can (and should) be maintained for other software developers: “The algorithm should be clean and understandable. [...] This makes it easy to extend its functionality. (E5)”

Thus, our expert participants viewed algorithms as strongly context dependent, that is, dependent on the task at hand: The problem at hand needs to be translated into manageable steps, frequently described as the algorithm’s logic. In other words, the decision-making process needs to be translated into a language that is appropriate to a computer. Thus, using certain data inputs reflects the solution of a problem. Data inputs were seen to be proxies for complex solutions that needed human interpretation. They were not directly understandable to the computer. E5 explained the logic in this way: “Logic is the abstract term for my problem-solving. You think of an approach of how to solve this problem [...] I ask myself, “okay, does the user have a certain preference? [...] The logic would be, for instance, to compare which movies I watched before, which ones I watched until the end, which ratings I gave the movie, how long I have read the message. Things like that. There are many things that flow into this. They reflect how much interest I had. ”

As there is a solution to the problem, the algorithmic system’s performance can be assessed through a benchmark. Therefore, many expert participants described the efficiency and runtime of an algorithm as important characteristics.

Finally, algorithms could also exist within algorithms, a perspective that was unique to the expert participants in our study. They were able to describe which processes occur within the black box of the algorithmic model, which was not surprising given their technical background. For example, the drawing of E11 showed an image recognition algorithm ([Figure 2](#)). The participant first described algorithmic systems and drew the general model of input, algorithmic model, output. Then, he explained the algorithmic model in detail by drawing a decision tree. Finally, each of the decisions in the tree was explained as comparisons of pixels within an image. While these elements were seen as one algorithm, still each element itself was also seen as an algorithm. This highlighted the nested aspects of algorithms.

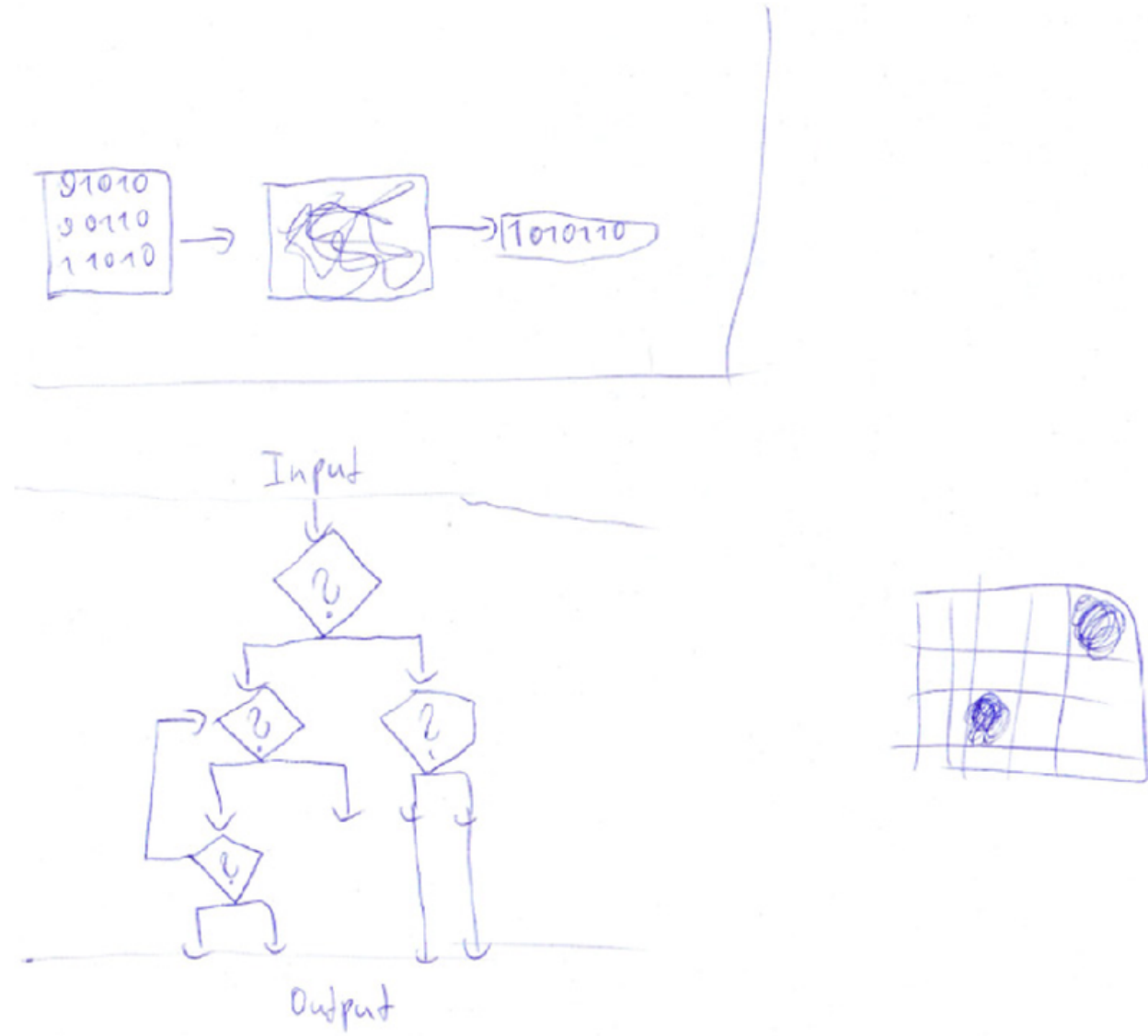


Figure 2
Drawing of E11 Illustrating Nested Algorithms

Lay Users' Theme 1: Algorithms as Autonomous Intelligence

According to the first author's interpretation, our lay participants associated the term algorithm strongly with (artificial) intelligence that acts independently and thinks on its own. Thus, the thematic analysis revealed that an algorithmic system's inner working was perceived as similar to a human mind's inner working. "I believe that algorithms are programmed to operate autonomously in the long run. And then just recognize behavior and patterns to predict what humans are interested in. In which direction this [the interest] might go. I think for this, there is artificial intelligence behind it. (L5) " " How they work? I believe it is like with us [humans]. If we get to

know somebody better, then we will have a more detailed impression of this person. It is like the famous rose-colored glasses if you have a crush on someone. Then you can have a wrong impression of this person that does not represent reality. And I think this is the same for an algorithm. It also learns more over time. This information can be wrong, and the impression of a person can be wrong. (L7) ”

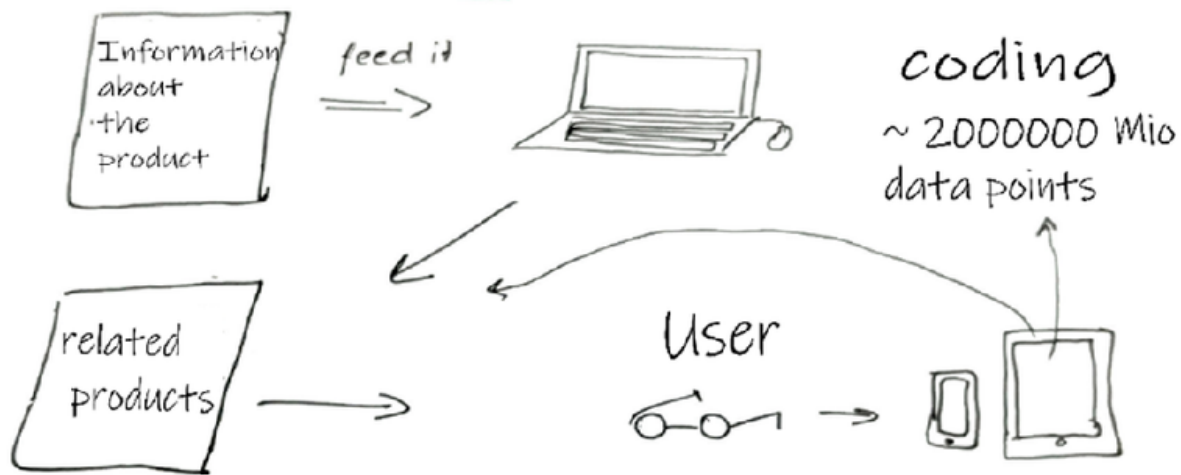
The analysis by the first author demonstrated that these intelligent algorithms exhibited human characteristics, such as being able to learn on their own, “Yes, of course, it can develop itself on its own through the PC.” (L4). It can follow “a human pattern” (L4). This was also evident in the drawing of L8, which compared algorithmic systems with a little child’s mind. Thus, algorithmic learning and reasoning were perceived to be equivalent to human learning and reasoning ([Figure 3](#)). The algorithmic autonomy and intelligence were seen as potentially dangerous. Lay participants perceived algorithmic systems as opaque, and thus, as uncontrollable and scary. “You always hear that this algorithm or the other algorithm did something for me. This is, for me, as a layperson, very non-transparent. You always have the feeling that algorithms act on their own. You, as a layperson, cannot influence how it works. This is also scary. (L4) ”

While the lay participants of our study perceived the algorithms as nontransparent, they believed that they themselves were quite transparent to the algorithm. Thus, transparency was seen as a one-way street: “I find it scary how much algorithms already know about me. For example, when you use social media, everything is tracked and saved and processed. I have been eleven years on Facebook. There must be a lot of information about me. (L1) ”

Thus, many of them pointed out that experts are necessary to understand algorithms. Algorithms were perceived as something complex, “Clearly, you need expert knowledge for this.” (L6). Given this black-box perception, it was unsurprising that many lay participants pointed out the societal impacts of algorithmic systems, such as data economy, specifically the sales of their personal data to third parties and exploitation of their data, but also the risks of political control through algorithmic systems. Thus, these systems were perceived as driven by economic interests and potentially harmful. “From [my] information, social media companies, but also others, make a lot of money. It is already known that much information gathered by algorithms are sold.” (L1)

L8 further explained that algorithmic news curation was not merely a commercial matter. He explained how filter bubbles posed intentional social control through

suggestion and manipulation.” You get chosen sources and media which reflect [your interest]. Thus, a bubble solution. So, you are only in the same area. For example, conspiracy theories of the coronavirus. It is an attempt for social and even political control. [...] Who is using this for which purpose? For the consumer, the user, it is non-transparent. And it is apparently the way it should be. ”



Data points: like teaching a child

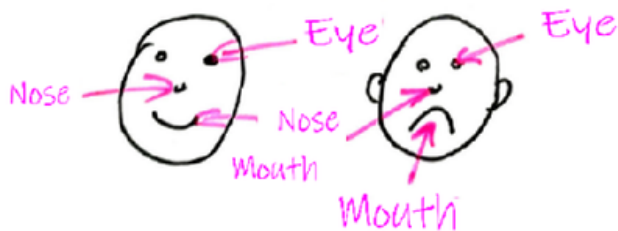
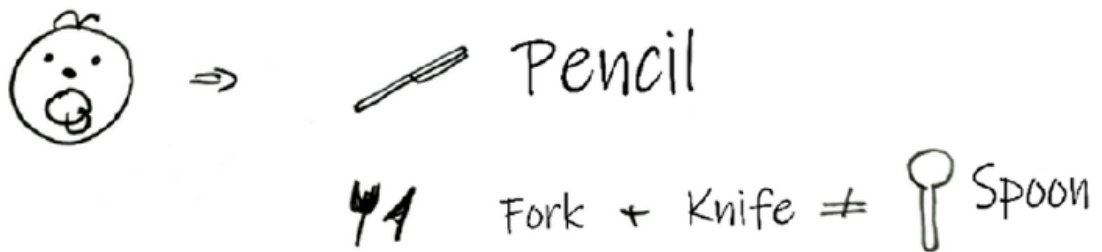


Figure 3

Drawing of L8 Illustrating Algorithmic Learning as Human Learning

Lay Users' Theme 2: Algorithms as Data Structuring

A second lay users' theme identified in our sample was the understanding of *algorithms as data structuring processes*. In this theme, algorithms were grounded in data or developed from data. "These are [data] resources that are gathered about me, and these are [data] resources about all others. From this, the algorithms are made. (L4) " " Algorithms are some sort of data sets, which are saved and automated so that I can see a lot of content which is interesting. Generally, [algorithms] are data sets that can recognize what is relevant for me. And later can present it [to me]. (L3) "

Thus, according to the first author's interpretation the term "algorithm" was perceived as equivalent to structuring the data to reveal certain patterns within the data, "You put an algorithm "on the data" to make something visible" (L6). As such, in this theme, algorithmic curation systems were often associated with data collection and analysis processes, including data filtering, saving, and systematization. This understanding was possibly driven by concepts from machine learning techniques in which features are extracted from a given training set.

In addition, other ideas discussed in this theme were related to user-collaborative filtering, as L4 described: "It [the algorithm] gets it [the data] from users who have the same interest. If my data is missing, then they will be complemented by other users." For him, this comparison of data was the central process for his understanding of algorithmic systems. Based on similar interests, users are grouped into profiles or schemas, for example, L4: "I think I will probably be put in a user group so that they can send me advertisements.."

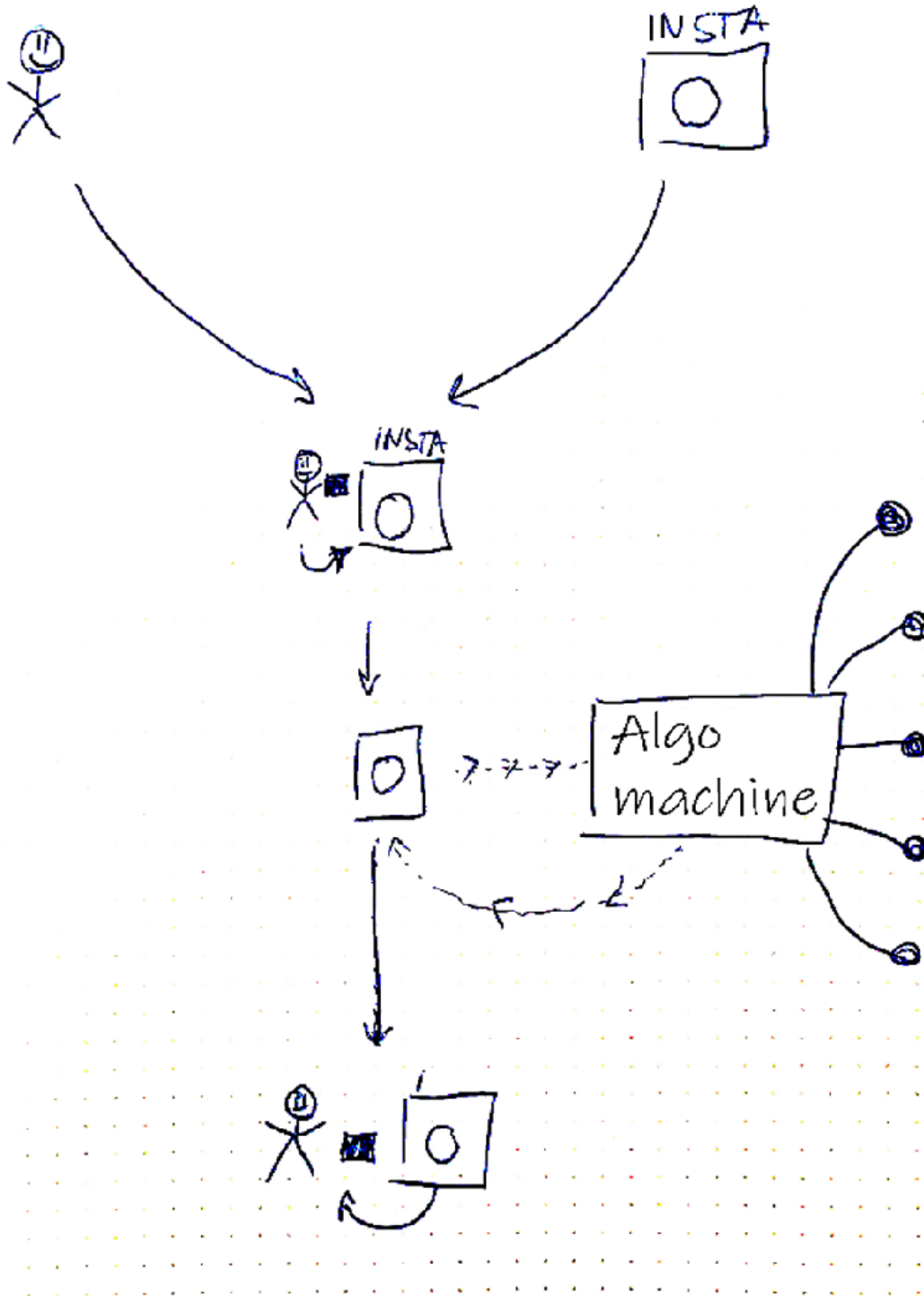


Figure 4

Drawing by L10 Expressing Overarching Structure Consisting of Three Elements: Data Input, Algorithmic Model, and Output

A Common Theme: Algorithms as Sequences of Actions

Reflexive thematic analysis by the first author revealed a common theme that was shared by our expert and lay users. This theme was characterized by two attributes: structural & systemic. Participants expressed a holistic view of the algorithmic systems, typically consisting of three elements: input, model, and output ([Figure 4](#)). Seventeen participants who expressed this type of view adhered to this basic structure. Three of them had nuances, including feedback loops or evaluation processes of the data. “I understand algorithms as a pre-defined and procedural set of sequences that can be programmed; they can fulfill a certain task. [...] Somebody has determined before how it works. (L10) ”

Our expert and lay participants viewed algorithms as a step-by-step workflow that one algorithm follows. Following this idea, both groups expressed that algorithms are built to pursue a certain goal. “Yes, so, whenever we want to build an algorithm, we, of course, have to think about what it should do and on which data it should run on.” (E7)

As such, the analysis by the first author showed that algorithms were viewed as neutral tools that software developers build. Understanding algorithms as a sequence of actions further implied that participants distinguished between theory and implementation. While the theoretical level was associated with mathematics, implementation of the algorithms implied more practical considerations, such as cost efficiency and customer satisfaction. “Algorithms are essentially mathematical functions. Just maths in beginner code. A bit exemplified and translated language that humans can understand. They are just very long mathematical functions.” (L11)“ There are algorithms that you cannot implement that well. [...] Sometimes, the theory helps you to decide. This is possible, and this is not possible. Then we have efficiency. When I see an algorithm, I ask [...] how well does it perform? How fast does it do it? Most of the time, we consider the pace. How long does the PC need to finish it? (E11)”

Our participants mentioned a variety of algorithmic tasks, including predictions, categorization, recognition of patterns. Participants were aware that these sequences work with data as they understood algorithms as sequences of actions. Thus, they were clearly distinguishing between algorithms and data. E12 pointed out that algorithms can be described as recipes, stating that algorithms were overestimated in the complexity. “When I used [the term] “algorithm,” then it’s usually among non-computer scientists. [...] I usually associate with something that is simpler than you think. Like a recipe which is executed by a computer. [...] I feel when talking to non-computer scientists that looking at an algorithm for the first time or understanding it might be

intimidating. [...] Yes, but if you look at the single pieces, step-by-step, then you can easily recognize that they are just a lot of small systems that were brought into one concept. ”

Summary

Reflexive thematic analysis revealed to the first author that experts and lay users in our sample shared an understanding of algorithms as sequences of actions that consist of input, algorithmic model, and output. Expert participants expressed a more abstract view on algorithms, understanding them as tools used for context-specific tasks. According to the first author's interpretation, this abstract understanding of algorithms allowed them to specify them for a different context. Hence, for the experts that were recruited in this study, algorithms are rather ways of thinking about a certain problem and translating them into a system to solve a certain task. In other words, they understood algorithms as a decision-making process which can be considered as an abstraction of the common theme ([Figure 5](#)).

In contrast to this, lay participants of this sample adhered to more visible and tangible elements, such as their user input and the output, and also physical devices (PC, server), neglecting the algorithmic model, that is, how the user data are specifically processed. Furthermore, they humanized algorithms ascribing them human-like features, such as intelligence. Thus, they lacked the awareness of the context dependency of algorithms. We argue that the lay users' understanding that was identified in this study can be considered as a simplification of the common theme.

Visit the web version of this article to view interactive content.

Explanatory drawings by lay users (original German; English translation)

Visit the web version of this article to view interactive content.

Explanatory drawings by experts (English translation)

Discussion

This study had two aims: (a) comparing the expert and lay users' understanding in order to identify the knowledge gap between these groups and (b) addressing implications for explanations that are informed by the lay user understanding. We first

discuss the knowledge gaps between the groups and then the implications for explanations.

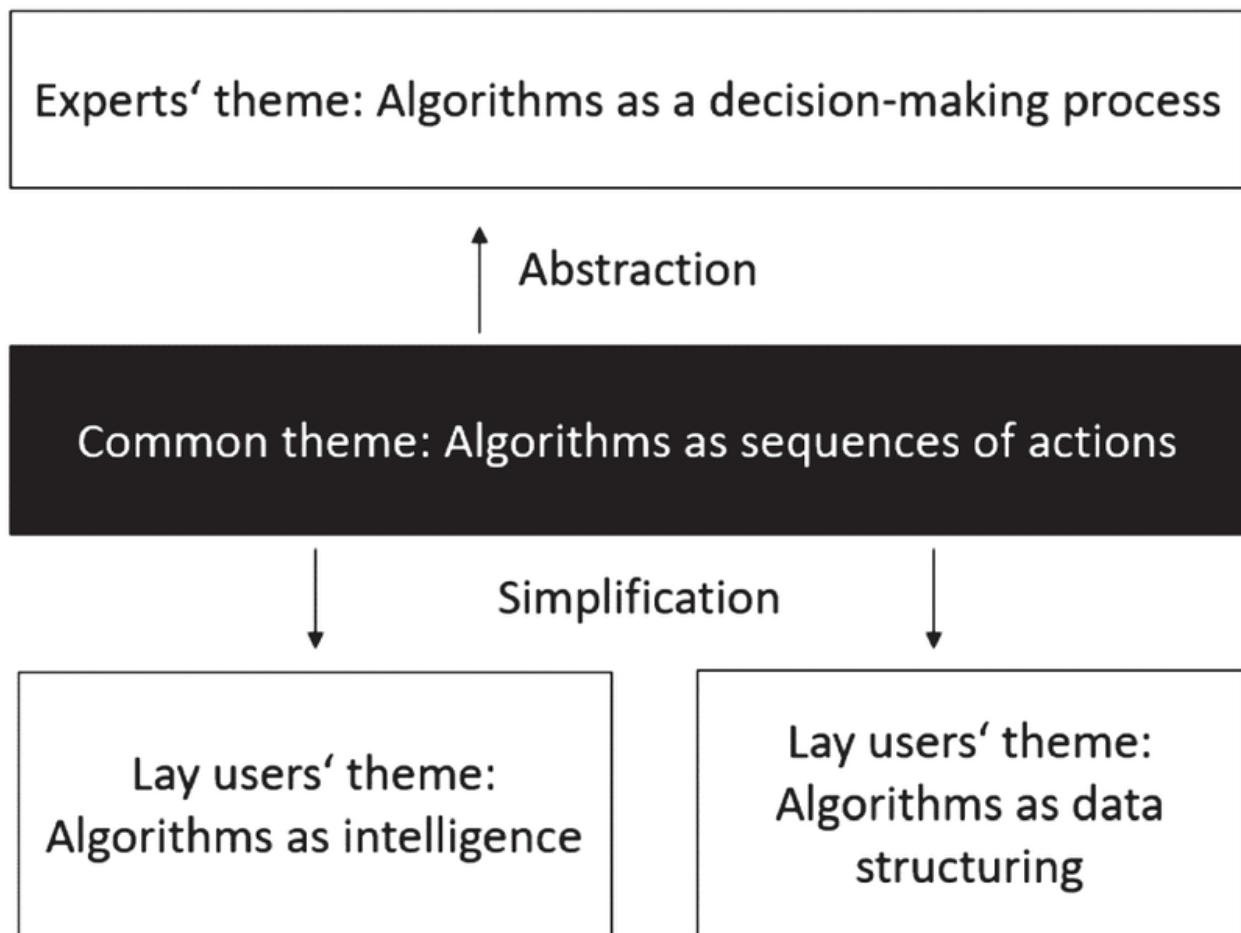


Figure 5
Experts' and Lay Users' Theme Can Be Interpreted as Abstraction and Simplification of the Common Theme

Both groups exhibited a structural understanding of algorithms consisting of the input, algorithmic model, and output. Nevertheless, within this understanding, the lay users in our sample did not exhibit an abstract understanding of the algorithmic model. Instead, they focused on visible and tangible elements of algorithms. This confirms previous results that have shown that experts' mental models are abstract and structural, while lay users' mental models tend to rely on visible elements ([Hmelo-Silver & Pfeffer, 2004](#); [Rouse & Morris, 1986](#)). In the context of algorithms, the lack of abstract understanding indicates that lay users had a narrower view of algorithmic curation. They were less aware of their context dependency than experts. In other words, lay users were not aware that algorithms work differently depending on the

task and goal in a given domain. This might be the reason why some lay users view them as being autonomous, intelligent, and data structuring.

The understanding of *algorithms as an autonomous intelligence* seemingly stems from media discussions on artificial intelligence (e.g., L3 specifically referred to the movie “I, Robot”), including societal consequences. Thus, the lay users’ themes go beyond the mental model and folk theories of how an algorithmic system works and a mere instrumental perception (i.e., viewing algorithmic curation as a tool). Our results add to the notion of algorithmic imaginary, which encompasses affective consequences of algorithms ([Bucher, 2017](#)). The lay users’ themes include aspects of algorithmic experiences such as data economy, privacy issues, and political control. Here, participants mostly experienced algorithms as a threat.

The theme of *algorithms as data structuring* focuses on machine learning concepts, that is, extracting patterns from data, neglecting that algorithmic curation systems also entail less sophisticated processes, such as data sorting. Like the other lay user theme, this theme incorporated tangible elements, that is, the data input and output. While the data input might not be visible, it is evident to many lay users that it stems from what they disclose to algorithmic systems, that is, what they like on social media or buy on e-commerce platforms. In this regard, our findings were similar to [Kang et al. \(2015\)](#). This form of understanding can be characterized as “simple and service-oriented” (p. 43).

The differences between experts and lay users of our study could be explained through the different motivations and purposes of each group to come in contact with algorithms. Algorithms can be described as *experience technologies*. As such, they are understood through use and interaction ([Cotter & Reisdorf, 2020](#)). Thus, most likely, our participants’ verbal and visual reflections of algorithms were influenced by their algorithms’ use. Here, in our sample, the experts mostly develop and study algorithms, and therefore, have a technical and task-oriented understanding of algorithms. Therefore, it seems that experts might develop awareness and understanding of the context dependency of algorithms. On the contrary, the majority of lay users experience algorithms when they use services. This means that lay users typically experience mostly the data input for algorithms, specifically what they are aware of disclosing to a system, and their results (e.g., personalized news feed). Outside this service, lay users primarily learn about algorithms in the media which include discussions about societal risks of algorithms. Thus, societal risks of algorithms were more frequently mentioned by lay users than experts.

Against the background of experience technologies, we would expect some differences in the lay users' understanding in the context of other algorithms, for instance, recommender systems (e.g., movie or music recommender). For them, we assume to find a similar service-oriented theme, in which algorithms are seen as data structuring entities.

All in all, our results show that the knowledge gap between experts and lay users mainly lies in the level of abstraction of their understanding. Thus, our experts were more aware of the context dependency of algorithms than our lay users. Lay users focused on more visible and tangible elements and viewed them as (a) an entity with human characteristics or (b) as a data-related process. Therefore, to address this fundamental knowledge gap, we suggest raising awareness of the context dependency of algorithms. We suggest that the task and goal of algorithms need to be emphasized. As a consequence, algorithms need to justify *why* certain user data are processed.

Creating User-Informed Explanations

Researchers have demanded an alignment of system developers' and experts' conceptual models and users' mental models to increase system transparency ([Eiband et al., 2018](#); [Norman, 1983](#)). Against this background, we conclude two implications for user-informed explanations.

First, explanations can rely on preexisting user knowledge incorporated in the common theme and could be composed of the overarching structure of the input, algorithmic model, and output. In this context, input explanations could entail the data used from the users, including personal data and external aggregated data from other platforms. To explain the algorithmic model, explanations could aim to frame the model as an abstraction, that is, explaining the algorithmic model as a decision-making process. This category is also known as how-explanations in the literature ([Friedrich & Zanker, 2011](#); [Rader et al., 2018](#)). Our results show that its context-dependency characterized experts' understanding of algorithms. This aspect was neglected in the lay user understanding of our sample. We assume that creating explanations that describe the step that a model takes to come to a certain output and emphasize the context dependency of the model might fill a knowledge gap between experts and lay users. However, we note that this needs to be addressed in future studies that test the impact of such an explanation on the user understanding. Finally, concerning algorithmic systems' output, system designers might want to keep in mind that lay users were aware of the output's visible aspects. If lay users should be informed about

output elements that are not directly visible (on the user interface), we suggest making this explicit to lay users.

Second, the lay user understanding in our study was characterized by humanization, that is, equalizing algorithmic systems' capabilities with human intelligence. Inherently human characteristics, such as autonomy and recognition, were ascribed to algorithmic systems leading to skeptical attitudes toward these systems' societal impact. At the same time, lay users' understanding did incorporate some technical knowledge on machine learning techniques (pattern recognition in training data). As some experts in our interviews have pointed out, this is an overestimation of algorithmic capabilities. The fundamental difference between algorithmic autonomy and reasoning and human autonomy and reasoning should be contrasted in explanations. Such explanations should underline the limitations of algorithmic systems.

Limitation and Future Work

While we acquired a diverse sample of experts from different computer science and related fields, our lay user sample was, on average, well educated and rather young. We were only able to reach participants who were proficient in using online tools. Thus, we assumed that our lay users' sample was, at least to a certain extent, interested in technology. For future work, we suggest investigating lay users who are unfamiliar with algorithmic systems, that is, do not use online tools regularly and might be rather skeptical toward them. Here, we could assume that these users might tend to oversimplify algorithms.

Additionally, we note that these qualitative results were analyzed and interpreted by the first author. They need to be replicated in larger, and possibly representative, studies to make generalizable assumptions about populations. For this, we suggest including measures of the abstraction level of knowledge, the knowledge of contextdependency, and the level of humanization to characterize the user understanding in large samples.

One follow-up question that this work raises is: "How effective are user-informed explanations in increasing the actual user understanding of algorithmic curation systems?." While this question was outside of our scope, we suggest addressing this question through experimental user studies using user-informed versus non-user-informed explanations as conditions. In this regard, it would be interesting to study the

effect of user-informed explanations on user experience measures (e.g., satisfaction, perceived ease of use), trust, and transparency.

Finally, the scope of our work was to characterize the user understanding of algorithmic systems. We did not investigate how this understanding is related to the interaction with these systems. Future research could investigate the role of understanding on behavior, such as privacy protection behavior, or other measures, such as trust or acceptance of systems' decision. Here, we speculate that lay users who adhere to algorithms' theme as intelligence might be more skeptical toward algorithmic systems.

Conclusion

Explanations to increase algorithmic transparency are predominantly informed by the system properties. Thus, they are not informed by the lay users' preexisting knowledge of the system. However, to ensure that explanations are understood accurately, it is necessary to tailor them to the users' knowledge. In this work, we compared experts' and lay users' understanding of algorithmic systems to reveal the knowledge gaps and misconceptions on algorithmic systems. While there is a common ground between these two groups, experts in our study exhibited an abstraction of this common theme while our sampled lay users simplified it. To overcome the knowledge gaps, a possible explanation could emphasize the context dependency of algorithmic systems and their application as a tool. Furthermore, the capabilities and limits of these systems should be highlighted to avoid overestimation of them.

Copyright © 2021 The Author(s)

Received January 8, 2021

Revision received June 12, 2021

Accepted June 29, 2021

Citations

1. Adadi, A., & Berrada, M. (2018). Peeking inside the black-box: A survey on explainable artificial intelligence (XAI). *IEEE Access: Practical Innovations, Open Solutions*, 6, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>
2. Craglia, M., Annoni, A., Benczur, P., Bertoldi, P., Delipetrev, P., De Prato, G., ... Vesnic Alujevic, L. (2018). *Artificial intelligence- a european perspective*. Retrieved

from <https://ec.europa.eu/jrc/publication/eur-scientific-and-technical-research-reports/artificial-intelligence-european-perspective->

3. Cotter, K., & Reisdorf, B. (2020). Algorithmic knowledge gaps: A new dimension of (digital) inequality. *International Journal of Communication*, 14. [↵](#)
4. Diakopoulos, N. (2015). Accountability in algorithmic decision-making. *Queue*, 13, 126-149. <https://doi.org/10.1145/2857274.2886105> [↵](#)
5. Tintarev, N., & Masthoff, J. (2007). A survey of explanations in recommender systems. In *2007 IEEE 23rd international conference on data engineering workshop* (pp. 801-810). The Institute of Electrical; Electronics Engineers. <https://doi.org/10.1109/ICDEW.2007.4401070> [↵](#)
6. Kunkel, J., Donkers, T., Michael, L., Barbu, C.-M., & Ziegler, J. (2019). Let me explain: Impact of personal and impersonal explanations on trust in recommender systems. In *Proceedings of the 2019 CHI conference on human factors in computing systems* (pp. 1-12). Association for Computing Machinery. <https://doi.org/10.1145/3290605.3300717> [↵](#)
7. Eiband, M., Schneider, H., Bilandzic, M., Fazekas-Con, J., Haug, M., & Hussmann, H. (2018). Bringing transparency design into practice. In *Proceedings of the 2018 conference on human information interaction & retrieval-IUI 18* (pp. 211-223). Association for Computing Machinery. <https://doi.org/10.1145/3172944.3172961> [↵](#)
8. Kang, R., Dabbish, L., Fruchter, N., & Kiesler, S. (2015). "My data just goes everywhere:" User mental models of the internet and implications for privacy and security. In *Proceedings of the eleventh symposium on usable privacy and security* (pp. 39-52). USENIX Association. [↵](#)
9. Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77-101. <https://doi.org/10.1191/1478088706qp063oa> [↵](#)
10. Braun, V., & Clarke, V. (2020). One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative Research in Psychology*, 18, 328-352. <https://doi.org/10.1080/14780887.2020.1769238> [↵](#)
11. Vaismoradi, M., Turunen, H., & Bondas, T. (2013). Content analysis and thematic analysis: Implications for conducting a qualitative descriptive study. *Nursing & Health Sciences*, 15, 398-405. <https://doi.org/10.1111/nhs.12048> [↵](#)

12. Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, 57, 227–254. <https://doi.org/10.1146/annurev.psych.57.102904.190100> [↵](#)
13. Došilović, F. K., Brčić, M., & Hlupić, N. (2018). Explainable artificial intelligence: A survey. *41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO)*, 0210–0215. [↵](#)
14. Edwards, L., & Veale, M. (2017). Slave to the algorithm? Why a right to an explanation is probably not the remedy you are looking for. *Duke Law & Technology Review*, 16, 18–84. <https://doi.org/10.2139/ssrn.2972855>. [↵](#)
15. Friedrich, G., & Zanker, M. (2011). A taxonomy for generating explanations in recommender systems. *AI Magazine*, 32, 90–98. <https://doi.org/10.1609/aimag.v32i3.2365> [↵](#)
16. Rader, E., Cotter, K., & Cho, J. (2018). Explanations as mechanisms for supporting algorithmic transparency. In *Proceedings of the 2018 CHI conference on human factors in computing systems-CHI '18* (pp. 1–13). Association for Computing Machinery. <https://doi.org/10.1145/3173574.3173677> [↵](#)
17. Balog, K., & Radlinski, F. (2020). Measuring recommendation explanation quality: The conflicting goals of explanations. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval* (pp. 329–338). Association for Computing Machinery. <https://doi.org/10.1145/3397271.3401032> [↵](#)
18. Miller, T. (2018). Explanation in artificial intelligence: Insights from the social sciences. *ArXiv:1706.07269 [Cs]*. [↵](#)
19. Chang, S., Harper, F. M., & Terveen, L. G. (2016). Crowd-based personalized natural language explanations for recommendations. In *Proceedings of the 10th ACM conference on recommender systems* (pp. 175–182). Association for Computing Machinery. <https://doi.org/10.1145/2959100.2959153> [↵](#)
20. Cai, C. J., Jongejan, J., & Holbrook, J. (2019). The effects of example-based explanations in a machine learning interface. In *Proceedings of the 24th international conference on intelligent user interfaces* (pp. 258–262). Association for Computing Machinery. <https://doi.org/10.1145/3301275.3302289> [↵](#)
21. Norman, D. A. (1983). Some observations on mental models. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 7–14). Psychology Press. [↵](#)

22. DeVito, M. A., Birnholtz, J., Hancock, J. T., French, M., & Liu, S. (2018). How people form folk theories of social media feeds and what it means for how we study self-presentation. In *Proceedings of the 2018 CHI conference on human factors in computing systems*, 120 (pp. 1–12). Association for Computing Machinery. <https://doi.org/10.1145/3173574.3173694> ↵
23. Eslami, M., Karahalios, K., Sandvig, C., Vaccaro, K., Rickman, A., Hamilton, K., & Kirlik, A. (2016). First i “like” it, then i hide it. In *Proceedings of the 2016 CHI conference on human factors in computing systems* (pp. 2382–2371). Association for Computing Machinery. <https://doi.org/10.1145/2858036.2858494> ↵
24. Gelman, S. A., & Legare, C. H. (2011). Concepts and folk theories. *Annual Review of Anthropology*, 40, 379–398. <https://doi.org/10.1146/annurev-anthro-081309-145822> ↵
25. Asgharpour, F., Liu, D., & Camp, L. J. (2007). Mental models of computer security risks. In S. Dietrich & R. Dhamija (Eds.), *Financial cryptography and data security* (pp. 367–377). Springer. https://doi.org/10.1007/978-3-540-77366-5_34 ↵
26. Hmelo-Silver, C. E., & Pfeffer, M. G. (2004). Comparing expert and novice understanding of a complex system from the perspective of structures, behaviors, and functions. *Cognitive Science*, 28, 127–138. https://doi.org/10.1207/s15516709cog2801_7 ↵
27. Jorgensen, Z., Chen, J., Gates, C. S., Li, N., Proctor, R. W., & Yu, T. (2015). Dimensions of risk in mobile applications: A user study. In *Proceedings of the 5th ACM conference on data and application security and privacy* (pp. 49–60). Association for Computing Machinery. <https://doi.org/10.1145/2699026.2699108> ↵
28. Renaud, K., Volkamer, M., & Renkema-Padmos, A. (2014). Why doesn't jane protect her privacy? In E. De Cristofaro & S. J. Murdoch (Eds.), *Privacy enhancing technologies* (Vol. 8555, pp. 244–262). Springer International Publishing. https://doi.org/10.1007/978-3-319-08506-7_13 ↵
29. DeVito, M. A., Gergle, D., & Birnholtz, J. (2017). “Algorithms ruin everything”: #RIPTwitter, folk theories, and resistance to algorithmic change in social media. In *Proceedings of the 2017 CHI conference on human factors in computing systems* (pp. 3163–3174). Association for Computing Machinery. <https://doi.org/10.1145/3025453.3025659> ↵

30. Morse, J. M. (2000). Determining sample size. *Qualitative Health Research, 10*, 3–5. <https://doi.org/10.1177/104973200129118183> ↵
31. Saldaña, J. (2013). *The coding manual for qualitative researchers* (2nd ed.). SAGE Publications. ↵
32. Rouse, W. B., & Morris, N. M. (1986). On looking into the black box: Prospects and limits in the search for mental models. *Psychological Bulletin, 100*, 349–363. <https://doi.org/10.1037/0033-2909.100.3.349> ↵
33. Bucher, T. (2017). The algorithmic imaginary: Exploring the ordinary affects of facebook algorithms. *Information Communication and Society, 20*, 30–44. <https://doi.org/10.1080/1369118X.2016.1154086> ↵

DuEPublico

Duisburg-Essen Publications online

UNIVERSITÄT
DUISBURG
ESSEN

Offen im Denken

ub | universitäts
bibliothek

This text is made available via DuEPublico, the institutional repository of the University of Duisburg-Essen. This version may eventually differ from another version distributed by a commercial publisher.

DOI: 10.1037/tmb0000045

URN: urn:nbn:de:hbz:465-20220727-153957-4

Publisher Website: <https://tmb.apaopen.org/pub/9oppvmbi/release/2>



This work may be used under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 License (CC BY-NC-ND 4.0).