Am I my data?

An investigation of how response-ability is created in *#2 BIAS*, a game on the manipulative power of algorithms

Elsbeth Hoefkens

BA Thesis Theatre and Dance

Tutor: Dr. Liesbeth Groot Nibbelink

Second reader: Dr. Jasper van Vught

Elsbeth Hoefkens (6143555)

Word amount: 8090

16-6-2022

ABSTRACT

Algorithms have a large influence on society. Although beneficial in many ways, they also cause social exclusion and ideological segregation. Literature about these consequences focuses mainly on individual experiences, although negative effects mostly become visible when looking at social groups. Where literature only writes about the consequences of algorithms, this performance provides a valuable live experience and a moment of reflection on this experience of the algorithms.

The performance *#2 BIAS* exposes the mechanisms of an algorithm that cause social exclusion. The focus of this research is on the way the audience is addressed. To analyse the modes of address I use the concepts of *theatricality* and *absorption* and the terms *discernable* and *integrated* relationships, originating from game theory. I use these concepts to analyse the notion of *response-ability*.

Following the strategies of an algorithm, *#2 BIAS* uses different modes of addressing the spectator. The concept of absorption is dominant, and, at times, there is theatricality. In the analysis of *#2 BIAS,* I show that due to the alternation between these modes of address, the spectator is invited to become aware of the game mechanics that correspond to the mechanism of an algorithm. Subsequently, a space for response-ability is created.

I analyse how the performance achieves this by using the dramaturgical analysis from Liesbeth Groot Nibbelink and Sigrid Merx, which entails a relational method based on three elements: context, spectator, and composition.

Even though at times the spectators are totally absorbed into the game, I will argue that due to accumulation of reflection moments, the audience is left with many things to think of and reflect on. By using the mechanism of an algorithm as dramaturgy, the performance thus exposes the way an algorithm works, lets the spectators feel the consequences of that in a group setting, and creates a place for the audience to reflect on both elements.

TABLE OF CONTENTS

1. INTRODUCTION

1.1 PHENOMENON AND *#2 BIAS*

After researcher, economist, and writer Sanne Blauw wrote an in-depth article for the Correspondent about the operation and function of algorithms this quote was her conclusion: "Algorithms are just a tool and whether we should be happy with them, depends on who makes the decisions. Even though algorithms seem abstract and automated, they are still human work. So above all, let us keep thinking about them." (My translation).[1]

More and more data are produced and collected these days. By sending emails, searching on Google, and posting tweets, we contribute to the massive amount of data called 'Big Data'. Once collected, the data is analysed by algorithms searching for patterns in the data and making a digital profile of a person. Judging by this digital version, predictions are made about, for example, someone's shopping preferences, online conduct, or criminal behaviour. These predictions can seem harmless but can be harmful. Algorithms thus play a big role in our (online) world.[2] In the previous quote Blauw emphasised the urgency of a critical attitude towards the use of algorithms. Researchers like Professor Safiya Umoja Noble, Eli Pariser, Cathy O'Neil, and Professor Cass Sunstein discussed the possible damaging effects of the predictions made by algorithms.[3] Pariser and Sunstein focused on the effect of political ideological segregation, arguing that the outcomes of algorithms used in social media create ideological segregation because of the lack of critical viewpoints.[4] Noble described the problem of how norms, values and prejudices of algorithm programmers influence the outcomes of algorithms, creating biased outcomes that can lead to social exclusion and racism.[5] Meanwhile, O'Neil zoomed out and described how Big Data increases inequality in general through the biased outcomes of algorithms that discriminate against people.[6] How exactly these biased outcomes come about will be discussed later.

[1] Sanne Blauw, "Wat is een algoritme?," *De Correspondent*, July 2, 2019, https://decorrespondent.nl/10306/wat-is-een-algoritme/149980270484-745de161. Original quotation: "Algoritmes zijn slechts een instrument en of we blij met ze moeten zijn, hangt af van wie de beslissingen neemt. Al lijken algoritmes abstract en geautomatiseerd, ze blijven mensenwerk. Dus laten we er vooral bij blijven nadenken."

[2] Thomas H. Cormen et al., *Introduction to Algorithms*, 4th ed. (Cambridge: MIT Press, 2022); NEMO, *Bits of You*, 2021, exhibition, 2021.

[3] Safiya Umoja Noble, *Algorithms of Oppression: How Search Engines Reinforce Racism* (New York: New York university press, 2018); Eli Pariser, *The Filter Bubble: What the Internet Is Hiding from You* (London: Viking, 2011), https://archive.org/details/filterbubblewhat0000pari_z3l4; Cathy O'Neil, *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy* (London: Penguin Books, 2016); Cass R. Sunstein, *#Republic: Divided Democracy in the Age of Social Media* (Princeton: Princeton University Press, 2018).

[4] Pariser, *The Filter Bubble*; Sunstein, *#Republic*.

[5] Noble, *Algorithms of Oppression*.

[6] O'Neil, *Weapons of Math Destruction*.

In the performance *#2 BIAS* (2021) by theatre collective playField, the audience is made aware of the potentially harmful role algorithms can play within a social group, with the theme of response-ability at its core.[7] It does so by using the mechanism of algorithms as a dramaturgical strategy. In doing so, the performance highlights the problem of biased algorithms which, among others, Professors in Psychology Annemarie Hiemstra and Isabelle Nevels point out in their article "Algoritmes leiden niet automatisch tot eerlijkere selectie" (Algorithms do not automatically lead to fairer selection).[8] In this article, they describe how algorithms, due to the subjective nature of data, produce perceptual bias within their predictions and therefore may cause exclusion based on prejudices, for instance in job application processes.[9]

The performance makes this social exclusion tangible and shows the influence of norms and values on the outcomes of the algorithms. By turning the spectators into the creators of the algorithms as well as the victims of its outcomes, the responsibility for their actions and consequences are directly and awkwardly felt by the audience. The performance thus centres on the theme of *response-ability*, a concept that Hans-Thies Lehmann succinctly describes as the created ability to react by taking action or actively reflecting.[10]

In analysing how response-ability is constructed in *#2 BIAS*, I focus on how the performance addresses the audience and thus how the performance invites the spectators to look at their actions and the way these actions are steered by an algorithm-based dramaturgy. When concentrating on modes of address, I use the concepts theatricality and absorption as defined by Professor of Theatre Studies Maaike Bleeker. I focus on the shifts between the two concepts to analyse the notion of response-ability. In addition, I use the terms discernable and integrated relationships used in game analysis as described by researchers and game designers Katie Salen and Eric Zimmerman.

[7] playField, *#2 BIAS*, 2021, theatre, 2021.

[8] Annemarie Hiemstra and Isabelle Nevels, "Algoritmes leiden niet automatisch tot eerlijkere selectie," *Sociale Vraagstukken*, December 3, 2018, 4.

[9] Byron Spice, "Questioning the Fairness of Targeting Ads Online," News Carnegie Mellon University, July 7, 2015, http://www.cmu.edu/news/stories/archives/2015/july/online-ads-research.html.

[10] Hans-Thies Lehmann, *Postdramatic Theatre*, trans. Karen Jürs-Munby (London: Routledge, 2006).

Figure 1. *#2 BIAS* in Heerlen, 2021. (Photo: Dani Silvia)

1.2 RELEVANCE

The debate of algorithms and how they function is characterised by many differing opinions. Numerous publications refer to the online use of algorithms and concern an individual experience of the effects of algorithms. By contrast, *#2 BIAS* focuses on the social and collective dynamics of exclusion processes and therefore offers a valuable additional perspective on existing literature. Only when looking at algorithms in the context of a group can the effect of social exclusion be explored. This research thus provides an opportunity to analyse the effects of algorithms on a group.

An additional advantage of a theatre performance as a case study is the emphasis on liveness and, in this case, the presence of other people, which offers the opportunity to not only make the effect of social exclusion visible but also tangible. Rather than reading about the harmful effects on a group, the audience members experience them directly. As a result, the spectators are enabled to actively reflect on that experience and can raise critical questions about the way we use algorithms.

1.3 RESEARCH QUESTION

To find out how the audience is addressed and how response-ability is created in the performance *#2 BIAS*, I will ask the question: *How do shifting modes of audience address in* #2 BIAS *construct response-ability?*

To answer this main question, I will first answer three sub-questions.

1. How do algorithms produce perceptual bias?

2. How is the notion of *response-ability* theorised by Lehmann and Ridout, and how do they establish a space for audience response?

3. What are the key shifts in modes of audience address *in #2 BIAS*?

By answering the last sub-question, I focus on changes in modes of addressing the spectators that occur in *#2 BIAS*. By analysing the key moments where these shifts appear, observations can be made about how these shifts construct response-ability. The concepts I use to analyse the key moments are the concepts theatricality and absorption, described by Bleeker, and the terms discernable and integrated relationships, explained by Salen and Zimmerman.

1.4 METHOD

When analysing the performance, I use the method of dramaturgical analysis following the approach of researchers Liesbeth Groot Nibbelink and Sigrid Merx in their article "Dramaturgical Analysis: A Relational Approach.".[11] This method consists of three relational elements: the *composition,* the *spectator,* and the *context* of the performance. These components interact with each other and together generate a meaning and experience of the performance.

In my analysis I focus, as already mentioned, on modes of addressing the spectator. Combining this approach with the composition of the performance and the social context of algorithms, observations can be made about the notion of response-ability.

To analyse the performance, I use a registration received by theatre collective playField on 9-3-2022. The registration is available on the platform *Vimeo* and takes 40 minutes.

---

[11] Liesbeth Groot Nibbelink and Sigrid Merx, "Dramaturgical Analysis: A Relational Approach," *FORUM+* 28, no. 3 (October 1, 2021): 4–16, https://doi.org/10.5117/FORUM2021.3.002.GROO.

Figure 2. *#2 BIAS* in Amsterdam, 2021. (Photo: Anna van Kooi)

Figure 3. *#2 BIAS* in Heerlen, 2021. (Photo: Dani Silvia)

2. THEORETICAL FRAMEWORK

2.1 ALGORITHMS AND SOCIAL EXCLUSION

The term algorithm is often associated with technology, but in a broader sense, an algorithm is a set of procedural steps which lead to a specific goal. This goal can be almost anything, for example cooking a meal. The total number of steps described in the recipe to cook a meal forms the algorithm.[12] The definition of an algorithm is thus very general but currently, the term algorithm is mostly used in the context of technology and the algorithms are way more complex than most cooking recipes. Today's algorithms are often used to calculate the probability of something, for example, how likely it is that someone will break the law or pay off their debt.[13] Algorithms use all sorts of data depending on their goal. For instance, an algorithm built to estimate the likelihood of criminal behaviour will use the person's address to see if there is a lot of poverty in the residential area, based on the assumption that there is a relation between income and criminality rate.

Algorithmic outcomes give a prediction of something and are often used to make choices. To stick to the previous example, when a person has a high score on the chance of becoming a criminal, the police can choose to act on that outcome and treat this person differently than people with a lower score. Here a possible problem, which I will elaborate on later, is the influence of perceptual bias which can affect the outcome. Mathematical Hannah Fry splits this choicemaking process up into four functions an algorithm can use: prioritise, classify, associate and filter.[14] In most cases, an algorithm combines multiple of these functions. For example, a police algorithm looks for associations within

---

[12] Max Vetzo and J. H. Gerards, "Algoritme-gedreven technologieën en grondrechten," *Computerrecht*, no. 1 (February 2019): 11.

[13] Roelf Anton Hoving, "Verdacht door een algoritme: Kan predictive policing leiden tot een redelijke verdenking" 7 (September 2019): 530–46.

[14] Hannah Fry, *Algoritmes aan de macht: hoe blijf je menselijk in een geautomatiseerde wereld?*, trans. Johannes Jonkers, 2nd ed. (Amsterdam: De Geus, 2018).

criminal behaviour and then classifies a person's risk level. Google uses the data collected to classify people, for example based on what hobbies they have, and then to prioritise and filter the information displayed.

Algorithms are thus always involved in decision processes and work on the basis of selection. Therefore, algorithms always produce a distribution pattern of inclusion and exclusion. The algorithm decides what someone sees, and what they do not see. This might sound like a problematic issue, but it does not have to be one. In many situations, it is very useful to have selecting algorithms that filter and order your information.[15] In medicine, for example, algorithms can help doctors interpret scans by identifying abnormalities in anatomical patterns, which usually makes the medical decisions better.[16] However, in other situations, algorithms can cause problems or discussions.

One of these discussions in the field of algorithms is about how the development of Big Data and algorithms influence our (online) daily life.[17] In this debate, researchers, including Pariser and Sunstein, show that the use of algorithms in social media increases political ideological segregation because people are mostly exposed to content conforming to their existing opinions and not to other critical viewpoints.[18] Companies like YouTube, for example, make money by keeping you on their website, therefore, they offer content that suits you. Other researchers, among which Eytan Bakshy, Solomon Messing, and Lada Adamic, argue that social media decreases ideological segregation because the increased offer of news articles online causes exposure to diverse ideas.[19]

A key issue related to algorithms is the perceptual bias within the creation process of an algorithm. The steps taken to reach the goal of an algorithm are made by humans, as well as the data collected to fuel the algorithm. This results in outcomes and thus choices that are biased because of people's ideas, blind spots, and prejudices. For example, an algorithm created to estimate if a person is more or less likely to become a criminal involves humans who decide which elements should be involved to assess if someone might be a criminal. To create this parameter, in some cases pre-existing data is used, but this data can also be biased.[20]

Predictive policing, for example, uses algorithms to predict whether a person is potentially a criminal or not. In America, as input for the algorithms, data was used that showed a relationship between African-American poor young men and crime, resulting in the police focusing on these

---

[15] Fry; Vetzo and Gerards, "Algoritme-gedreven technologieën en grondrechten," 11.

[16] Eric J. Topol and Abraham Verghese, *Deep medicine: how artificial intelligence can make healthcare human again*, First edition (New York: Basic Books, 2019).

[17] Seth Flaxman, Sharad Goel, and Justin M. Rao, "Filter Bubbles, Echo Chambers, and Online News Consumption," *Public Opinion Quarterly* 80, no. S1 (2016): 298–320, https://doi.org/10.1093/poq/nfw006.

[18] Pariser, *The Filter Bubble*; Sunstein, *#Republic*.

[19] Eytan Bakshy, Solomon Messing, and Lada A. Adamic, "Exposure to Ideologically Diverse News and Opinion on Facebook," *Science* 348, no. 6239 (June 5, 2015): 1130–32, https://doi.org/10.1126/science.aaa1160.

[20] Hoving, "Verdacht door een algoritme," 84–89.

characteristics.[21] This form of ethnic profiling does not only harm innocent people but also likely arose in part because of discrimination by the police. Professor of Information Studies Safiya Umoja Noble writes about this form of racism as a problematic effect caused by subjective algorithms.[22] Similar to Hiemstra and Nevels, Noble observes that algorithms, because of the involvement of people's prejudices, create perceptual bias within their outcomes.[23]

The effects and use of algorithms thus vary strongly, from being helpful by making processes easier and better, to being harmful because of the perceptual bias within the outcomes. The outcomes of the decision-making process which involve inclusion and exclusion can thus be innocent, but also can lead to extreme versions of social exclusion and ideological segregation.

2.2 RESPONSE-ABILITY

The performance *#2 BIAS* uses the bias of the spectators in the creation of the algorithm, but also exposes this bias and its consequences. As a result, the notion of response-ability is created. Professor of Theatre Studies Hans-Thies Lehmann introduces this term in his book *Postdramatic Theatre*, in which he focuses on new theatre forms that have developed since 1960.[24] Lehmann mentions the notion of response-ability when considering the political influence of postdramatic theatre in a strongly mediatised society. He argues that the political engagement within theatre is not reliant on the subjects, but on the forms of perception.[25] Theatre can "intervene at the level of 'perception', by activating a capacity to respond (response-ability)".[26] Lehmann defines the word 'respond' here not only as action, but also as active reflection on what is seen. What this intervention entails and how this ability is achieved comprises various elements.

Firstly, postdramatic theatre, as Lehmann describes, is characterised by showing different perspectives and inviting the audience to assign their own meaning to what they see. In doing so, theatre may activate the spectator's thinking about perception and question their own perception, ultimately leading to a change in behaviour.

Secondly, in a world where the media shapes our perception of things, Lehmann states that, because of the massive amount of information available and the unclear separation between what is

---

[21] Bradley R. Entner Wright and C. Wesley Younts, "Reconsidering the Relationship between Race and Crime: Positive and Negative Predictors of Crime among African American Youth," *Journal of Research in Crime and Delinquency* 46, no. 3 (August 1, 2009): 327–52, https://doi.org/10.1177/0022427809335170; Blauw, "Wat is een algoritme?"

[22] Noble, *Algorithms of Oppression*.

[23] Hiemstra and Nevels, "Algoritmes leiden niet automatisch tot eerlijkere selectie."

[24] Lehmann, *Postdramatic Theatre*.

[25] Lehmann, 184.

[26] Nicholas Peter Ridout, *Theatre & Ethics*, Theatre & (London: Palgrave Macmillan, 2009), 57.

real and what is not, we lose connection with the source from which these messages originate. Consequently, the senders of the information are freed from the link with their message.[27] Meanwhile, by becoming disconnected from the source, the receivers lose the sense of being responsible for what they perceive. An example to clarify this 'receiver responsibility' is that, when people are on the beach and someone appears to be drowning, witnesses of the situation are likely to feel a sense of responsibility to take action. Lehmann argues that this disconnection from our perceptions, which results in a lack of responsibility, is a contemporary problem in which theatre can play an important role. Because of the liveness of theatre, theatre can create a strong connection between the viewer and what is viewed. Therefore, it may generate a feeling of responsibility towards what is being seen.

Professor Nicholas Ridout continues to explain his perspective on response-ability in his book *Theatre& Ethics*.[28] Ridout positions this term in the field of ethics within theatre, where he connects response-ability to the fundamental question of how humans may or should act.[29]

Ridout continues with the same broad interpretation of the word 'response' as Lehmann. In addition to Lehmann's analysis Ridout describes the way that "the global media typically preclude any response".[30] Theatre, however, "makes the possibility of response central to the way it functions" because spectators and performers are in the same space.[31] This makes the spectators actively aware of being part of the experience, as in the beach example, and do not become "a passive recipient of media".[32] Here, the sense of responsibility returns again.

Because theatre creates a strong connection with the source of the experience and makes the spectators part of an experience, Ridout argues that theatre invites the audience to connect what they see on stage to their personal life. He mentions an essay by Peggy Phelan, who writes about a performance by Marina Abramović where the audience becomes a witness: the spectator is not a voyeur but is acknowledged by the performer.[33] The performance, *The House with the Ocean View* (2002), is an experiment of Abramović living in three minimally furnished raised 'rooms' in a gallery without food, for twelve days straight. The audience could look at her and Abramović could look back at the audience. Here, Phelan emphasises the simultaneous presence of performer and spectator but indicates that this goes beyond that immediate personal situation.[34]

---

[27] This makes it easier to post things anonymously on social media, for example.
[28] Ridout, *Theatre & Ethics*.
[29] Ridout, 1.
[30] Ridout, 57.
[31] Ridout, 57.
[32] Ridout, 58.
[33] Peggy Phelan, "Marina Abramovic: Witnessing Shadows," *Theatre Journal* 56, no. 4 (2004): 569–77, https://doi.org/10.1353/tj.2004.0178.
[34] Ridout, *Theatre & Ethics*, 61.

Important here, is the role of a face-to-face encounter, which Phelan, referring to the French philosopher Emmanuel Levinas, describes as a crucial element of theatre because it creates a "general sense of connection to one another that exceeds simple geophysical, ideological, or cultural proximity."[35] Referring to the ethical question of how humans can or should act, the result here is that the spectator as a human being can be confronted with this question without all the things that emphasise the differences between people like cultural or ideological elements.

Ridout further indicates that Phelan's analysis is about what this face-to-face encounter evokes for the audience in terms of thoughts and connections beyond their personal lives. Ridout writes: "an ethical 'reawakening' might help us think and feel about those others we only ever encounter as images amid the media saturation".[36] In the performance of Abramović, the audience might think about how fragile the human body is and about their role in helping others that suffer. Consequently, they might think about giving Abramović food and whether they can and should interfere with this work of art in this way.

In conclusion, by "re-situating" the experience from the media to the theatre the connection between the source and the perception is recovered and the opportunity to respond is replaced.[37] Both result in active involvement of the spectator and therefore create a feeling of responsibility. Moreover, the role of face-to-face contact can make the spectator a witness and can create a sense of connection from person to person without the elements that separate us. As a result, people are likely to connect to what is seen on stage and to reflect on that experience. Here, the ability to respond is created.

2.3 THEATRICALITY AND ABSORPTION

In line with the used method, I focus on the spectator by using the concepts of theatricality and absorption from Bleeker to analyse how the spectator is addressed in the performance.[38] A characteristic of this performance is that it not only uses but also exposes the mechanism of algorithms. The audience is put in charge of creating the algorithm and confronted by the results of this. Subsequently, spectators are made aware of the construction of the performance. Bleeker terms this strategy of exposing the way the performance mediates in addressing the audience *theatricality*.[39] In such situations, the way the performance addresses the audience is visible. As a result, the

---

[35] Phelan, "Marina Abramovic," 577.
[36] Ridout, *Theatre & Ethics*, 61.
[37] Ridout, 58.
[38] Maaike Bleeker, *Thinking through Theatre and Performance*, ed. Adrian Kear, Joe Kelleher, and Heike Roms (London; New York: Bloomsbury Publishing, 2019).
[39] Bleeker; Lehmann, *Postdramatic Theatre*.

spectators are aware of their presence and the point of view offered by the performance. This can be done, for example, by suddenly addressing the audience directly or by disrupting the performance by adding an alienating element in a realistic situation. The opposite effect is what Bleeker calls *absorption*, which she defines as an invisible mode of communicating with the audience.[40] The way the audience is addressed by the performance is concealed, and so the spectators are unaware of the theatrical construction and of the fact that they are being addressed in a specific way. Consequently, "the seer takes up the position or point of view presented to him or her and does so without giving it a second thought."[41]

Theatricality and absorption strongly connect to the context of the seer or spectator because both concepts are effects on the viewer and thus occur together with the viewer. For example, absorption may arise when the way of addressing fits the audience's historically and culturally defined expectations because, in this case, the spectators experience no other perspective than their own. Bleeker writes about this: "Absorption does not refer to an a-historical quality of a work but is the effect of the interaction between a work of art produced at a particular time and place and a historically and culturally specific viewer."[42] Thus, the spectator is always already culturally and historically framed, which affects how the audience perceives the experienced performance.

2.4 DISCERNABLE AND INTEGRATED GAME DYNAMICS

In addition to absorption and theatricality, I use the terms discernable and integrated relationships to analyse the modes of address in *#2 BIAS*. These terms are defined by researchers and game designers Katie Salen and Eric Zimmerman in their book *Rules of Play*, wherein they discuss the fundamentals of game design.[43] In the analysis, I show that the performance can be understood as a game in which the theatre conventions are the rules, and the spectators are the players because of their active participation and significant influence in the performance. The main goal Salen and Zimmerman mention in designing a game is to create "meaningful play".[44] This meaningful play "emerges from the relationship between player action and system outcome".[45] How meaningful the game is depends on to what extent this relationship is "both *discernable* and *integrated* into the larger context of the

---

[40] Maaike Bleeker, *Visuality in the Theatre: The Locus of Looking* (London: Palgrave Macmillan UK, 2008), https://doi.org/10.1057/9780230583368.

[41] Bleeker, 33.

[42] Bleeker, 22.

[43] Katie Salen and Eric Zimmerman, *Rules of Play: Game Design Fundamentals* (Cambridge, Mass: MIT Press, 2004).

[44] Salen and Zimmerman, 37.

[45] Salen and Zimmerman, 34.

game".[46] By a *discernable* relationship, Salen and Zimmerman mean "that a player can perceive the immediate outcome of an action".[47] For example, if you play the game Mario Kart and you bump into someone, there must be a noticeable reaction from your action: a sound, a change of direction, or both. If the kart you bumped into kept its course, then there is no meaning, and thus less fun, in bumping against other karts. Within an *integrated* relationship, the outcome of the action is "integrated into the larger context of the game".[48] A good example here is the game chess because all your actions influence the rest of the game, so every move is integrated into the larger context and the outcome of the game.

Both terms refer to the effect of the actions from the choices provided by the game. To analyse how these choices are designed and therefore create a specific effect, Salen and Zimmerman created five questions that reveal the construction of the choice, which I use to analyse which effect the design of the choices in *#2 BIAS* creates. [49] These questions are as follows:

1. What happened before the player was given the choice?
2. How is the possibility of choice conveyed to the player?
3. How did the player make the choice?
4. What is the result of the choice? How will it affect future choices?
5. How is the result of the choice conveyed to the player?

---

[46] Salen and Zimmerman, 34.
[47] Salen and Zimmerman, 37.
[48] Salen and Zimmerman, 35.
[49] Salen and Zimmerman, 63–65.

3. ANALYSIS

#2 BIAS is performed by theatre company playField and uses the format of a game to address the topic of algorithms. The space of the performance is comparable to a lecture hall: in the front, there is a blackboard and chairs opposite it in the shape of a grandstand. Before entering the performance space, data from the audience is already collected when they arrive at the building entrance. From each spectator, the name is written down and a photo is taken. Subsequently, they are given an identity number by a sticker on their clothes. The numbers given to the audience members correspond with the numbers noted on white balloons adjusted to the chairs in the auditorium. On the blackboard, two lists of opposite social categories have been written and when starting the performance, the host adds her name to the board while welcoming the spectators and explaining to them that she is their moderator for this evening. The audience sits close to each other and the board. Due to the moderating host, the numbers given to the spectators corresponding to seats, which make one suspect that the audience must participate, at the start the performance has the setting of a social game.

In my analysis of the performance *#2 BIAS*, I differentiate three different phases corresponding to three key moments in the performance. I analyse these moments in chronological order. Phase one is characterised by the dominance of the strategy of absorption. Here the audience members are, most of the time, unaware of the structure of the performance which entails them building an algorithm. In the second phase, the spectators are still absorbed in the performance but are also at some moments made aware of the mechanism of the performance. This phase also includes the first moment of confrontation with the outcomes of the audience's actions. However, the analysis shows that the performance is designed in a manner that does not leave much room for response-ability yet. The third phase does leave room for evaluation by absorbing the audience first and then exposing the construction of the performance and the algorithm completely. The result at the end is the confrontation of the audience with its behaviour during the performance. This confrontation is an accumulation of moments of exposure built up through the performance that create response-ability.

3.1 PHASE ONE

At the beginning of the performance, the host introduces herself with the text "I am Carine, and I will guide you through this performance". She starts to mention facts about herself, for example, "I am a woman, 1.70 metres long and my weight is 64 kilograms". After numerous facts, she distinguishes these facts from the character traits she would use to describe herself like shy, clumsy, or friendly. While the host is telling this, she makes friendly eye contact with each individual, which she continues to do during the entire performance. The performance continues with a question for the audience,

they are invited as a group to make social categories based on labels, they normally give people they have just met, like shy, outgoing, sexy, sloppy, or friendly and unfriendly. In making the list, they can use the categories already written down on the board, but they are free to suggest new ones. For each suggested category, the audience is addressed as a group and asked if they agree with the category, after which the categories are written down on the board. In this category-making process, the host focuses on the agency and will of the group by using sentences such as "the group decides" and "does the group agree with these categories?". This creates a collective responsibility because, if you do not agree, you have the opportunity to indicate this.

Figure 4. *#2 BIAS* in Heerlen, 2021. (Photo: Dani Silvia)

Absorption/theatricality

The host invites the audience to come up with categories in a lighthearted and playful manner. She used sentences like "I would like for you as a group to make a list of your own". She explains clearly what the task is and checks if the spectators understand, but what she leaves out is the reason the social categories are made. Since the real construction has been hidden, and the audience is not aware of it, this scene can be understood as one in which the strategy of absorption is in operation.

Discernable/integrated game dynamics

The outcomes of the audience's actions are, as the analysis will show later, integrated into the larger context. Besides that, the outcomes of the spectator's actions become directly visible – the categories on the board – therefore, the relation between action and outcome is discernable. This discernability ensures that the players know what is happening when they act. What happens here, however, is that despite the discernability, players do not know the long-term consequences of their actions.

This mechanism of apparent clarity matches the way an algorithm works while using algorithm-driven platforms. An algorithm works invisibly, which is why the user does not know what the influence of the algorithm is. There are no warning pop-ups that inform the users their displayed information is filtered, excluding other types of information and ideas that do not match their profile. The performance thus approaches the audience here as the maker of the algorithm. Furthermore, the construction of this moment corresponds with how the parameters of a real algorithm are created, namely by humans with subjective ideas about which social categories are important.[50]

3.2 PHASE TWO

What the audience does not know when making the categories, is that they are making the parameters for an algorithm. After the audience has made social categories, they are asked to look at the individual number they have been given and take a seat at the corresponding balloon in the auditorium. Here they find a headset with a microphone. Half of the seats are also equipped with a small camera in front of them and the other half of the seats have a small touchscreen device. When they arrive at their places, the audience is invited to put on the headset and start short conversations with another person from the audience who is connected to their headset. The conversations take place at a distance, so both persons cannot see each other live, only the persons asking the questions see their conversation partner on their touchscreen device. In the conversations, one audience member asks pre-set questions found on the device, and the other answers them. After the short conversation, before changing partners, the touch screen device asks the spectators who asked the questions to label the person answering according to the social categories collected earlier. For example, the audience member needs to choose between shy or outgoing and friendly and unfriendly, and so on.

Beforehand, the host says, "Now it is time to get to know each other" and explains the structure of the conversations but leaves out the labelling part. After the third question round, the audience member who asked the questions needs to choose via the touchscreen display the labels that fit a 'trustworthy' person. Once this is done, the audience is asked to switch positions and have three more conversations. Now, the roles of asking and answering questions are reversed. During the six conversations, the blackboard is filled with all the data the technical crew of the performance have heard while listening in on the conversations. When the conversations are done, the host emphasises this collection of data by saying: "We are doing great as a group. Look at how much input we have already generated.".

---

[50] Hiemstra and Nevels, "Algoritmes leiden niet automatisch tot eerlijkere selectie."

The performance continues with the outcomes of the labelling process. Based on the categories the audience gave to each other and the categories the spectators linked to the term 'trustworthy', the audience members are clustered in the groups based on the amount of 'tags' signalling trustworthiness. First, the host lists some numbers referring to the individual numbers of spectators and explains that these people, according to the group, are the most trustworthy. She asks for a big applause while inviting the people mentioned to come down to stand on the stage as a group. This group is now called the "trust engineers". Secondly, the host again lists some numbers and labels them as "semi trust engineers". While inviting them onto the stage, the host asks for applause, but tells the audience that there is room for improvement for these people. The last group, which the host labels as "minor trust engineers", are welcomed to the floor without the host asking for applause and with the commentary that these people have a lot of room for improvement. As a result, after a moment of silence, the audience starts clapping hesitantly of its own accord.

Until this moment, the host related everything that happened in the performance to the decisions of the group by saying things like "according to the group" and "the group finds". After the split-up of the group, she mentions that they are now going to make an algorithm. What she does not mention is that the already collected data is, in fact, already a part of the algorithm.

Figure 5. *#2 BIAS* in Heerlen, 2021. (Photo: Dani Silvia)

Figure 6. *#2 BIAS* in Heerlen, 2021. (Photo: Dani Silvia)

Absorption/theatricality

Again, in this second moment, the task given to the audience is clear: ask or answer the pre-set questions. Here, absorption arises because the audience is unaware of the underlying mechanism of the performance and is not questioning the intentions behind their tasks. Because of the host saying, "how much input they already generated" and by presenting the results of the labelling process, the impact of the spectator's actions become clear as well as the possible harmful effects of an algorithm.

Namely, calling people unreliable and treating people unequally because of the labels other people have given them. Here the mode of audience address shifts towards theatricality, because the confrontation is a disruption of the playful, friendly atmosphere of the performance and reveals the construction of the performance. This change from absorption to theatricality confronts the audience and has the potential to create room for response-ability, because the audience is aware of the construction and can therefore reflect on its functioning or even take physical action. The moment where the audience starts to clap on its own initiative, shows this room where the spectators reflect on the situation and act in the situation. After this moment of clapping, the audience continues without hesitating. Therefore, it is interesting to ask why the spectators do not intervene and why thus the space for response-ability is not sufficient to create a bigger response.

Similar to the first moment, the host invites the spectators politely to do what she explains to them. She uses phrases such as "please answer the question" and "just be yourself", which often assume that the audience will participate. As a result, the spectators will have to actively oppose the invitation if they do not wish to participate. Besides that, the action that the spectators have to do when they participate is made simple, which makes it easy to act. The audience acts by answering or asking pre-set questions and touching the chosen categories on a touchscreen display. This last action is also anonymous, which makes it less hard to label someone.

Discernable/integrated game dynamics

Another element that can explain why the audience continues to act is because the performance integrates the results of the choices made by the spectators. The social categories made at the beginning are now the options for the second moment of acting. The result is thus the foundation for the next choice which, according to Salen and Zimmerman, creates a meaningful game for the player because their actions are integrated into the game. The effect on future choices is therefore a positive one, which includes a willingness to act, because they feel that their choices matter.

Looking at *how* the performance integrates the outcomes of the actions of the audience is also useful to analyse why the audience continues. To start, the audience does not know what the results of their actions will be used for, so the spectators cannot oversee their impact. In some video games, for example, as a player, you know if you press A or B what the results will be, and you probably also see the direct outcome. The consequence of this is that, as a player, you get the direct feedback of your actions. In *#2 BIAS*, this feedback loop has a delay. The spectators have conversations first, then label each other, then relate social categories to trust, and then it all repeats itself. After more than 18 minutes, the outcomes are presented, and the painful feedback reaches the audience. The relation

between the outcome and the action is thus not discernable. This lack of immediate response likely reduces the emotional impact of the outcome, and therefore may reduce the feeling of responsibility.

In this phase, the audience is again responsible for making the algorithm but is also a victim of its outcomes. Splitting up groups into more or less trustworthy creates a first moment in which the group feels the negative effect of the algorithm they themselves created. Here, the influence of people's subjectivity on the outcomes of an algorithm becomes visible because the used group labels are not made by an 'objective' computer but by the spectators themselves when they are asked to define reliable characteristics. The difference with a real algorithm is that now the people who make the algorithm are visible, with as a result that the subjective human influence is visible.

Another difference from a real online algorithm is that, in this performance, the outcomes are felt within a group and not isolated behind a computer screen. The painful effects are therefore awkwardly felt and become visible. This is an important element because algorithms thrive by having invisible or abstract effects. After all, when they become visible, they can be analysed critically and changed. When people have felt that the results of algorithms can have harmful exclusionary effects, they will become more critical about them.

3.3 PHASE THREE

The third and last moment the spectators are asked to act is near the end of the performance. Here the host raises questions about the composition and ideology of the group. For example, "Does the group need a leader?", "Do we want sexual tension in our group?" and "Is a diverse group important?". The members of the audience need to answer these questions from their idea of an ideal group. They do so by choosing to walk to the left or right side of the stage, which corresponds with either 'yes' or 'no'. The host explains the task clearly and tells the audience that they are going to improve the algorithm they created so far. After six questions the host says, "We have combined our data with our decisions. Our algorithm is working perfectly now, so let's put it in use. Let's make our group better".

The host continues and starts to read some facts which according to the created algorithm define the group. Facts like "if you have leadership qualities, you cannot be part of this group" and "if you are a fuckable person, you cannot be part of this group". These facts are on a set of cards that she was given by one of the technicians. After some facts, one of the technicians comes on stage and starts to name a number corresponding to one of the audience members and the reason this person needs to leave the group. For example, "Number two, you have to leave the group because we don't need a leader" and "Number one, you have to leave the group because you are the most fuckable". This element repeats four times, each time with a different reason. During this part of the performance,

21

the spectators laugh and behave as if it is not a big deal when spectators are asked to leave the group. The response of the audience, however, changes quickly when the host is finished with the outcomes of the algorithm and asks the excluded spectators to leave the performance for real. The audience becomes quiet, looks awkwardly at each other and after the performance ends with the host saying "congratulations, this is our group", they applaud hesitantly and leave the performance quickly.

Figure 7. *#2 BIAS* in Heerlen, 2021.

Figure 8. *#2 BIAS* in Heerlen, 2021.

Absorption/theatricality

Before the audience's task of answering the questions, the host explains what they are going to do and to what end – making the algorithm better – but not what the concrete effects will be when they answer the given questions. Thus, the host is again seemingly open about the construction of the game, but in fact leaves out important information. The real construction is hidden, and the audience joins in without a second thought; the strategy of absorption is thus in use. While answering the questions asked by the host, the audience is once more absorbed in their task, not thinking about the possible impact of their answers. This absorption changes into theatricality when the host says, "Let's put the created algorithm to work" and as a result, people are removed from the performance. Here, the spectators are reminded that the performance was about creating an algorithm, and they now experience exactly what that means. Because this moment of exposing the mechanism of the performance is not followed by an element that absorbs the spectators again, this moment of theatricality has much more impact in the sense that the construction is permanently exposed, and the audience cannot escape this reality.

Discernable/integrated game dynamics

Again, here the consequences of the audience's actions are integrated because their answers are used to finish the algorithm. However, the audience does not know what the effects of their actions will be.

As opposed to phase two, the result of the public's actions is integrated differently, because now the result harms not a group of people but individuals. Because the outcome of the action is again not directly visible, this relation is, similar to phase two, not discernable. Unlike the other moments, after this confrontation, there is no new task that makes the audience forget what they did and failed to do. They have room for reflection as the performance ends with an awkward situation and perplexed spectators, leaving the audience to realise what has happened.

In this third moment, the influence of humans with their subjective ideas and prejudices on the outcome of algorithms is most obvious. The spectators are openly asked to give their opinion about the group. Nothing is hidden or anonymous anymore, and the effects of their choices are therefore felt by each individual spectator. This makes the effect even stronger when individuals are removed from the performance because of the results of the questions about the group.

3.4 HOW AN EXPERIENCE OF RESPONSE-ABILITY IS CONSTRUCTED

Combining the insights of the previous analysis, I analyse how *#2 BIAS* constructs an experience of response-ability.

In the analysis of the different moments, it becomes clear that the performance switches between moments of theatricality, where the audience is made aware of the construction of the performance, and moments of absorption where the spectators forget they are addressed in a specific way. The moments of theatricality make room for moments of reflection because the mechanism of the performance is highlighted, reminding the audience what they are part of and thus creating the notion of response-ability. Analysing these moments, people indeed look confronted with the outcomes and have a small reaction, but do not hesitate with the following task. The performance gives too little room for a profound reflection at that moment of theatricality because it continues almost directly with a new task. In doing so, the spectators are tempted to continue as well.

In addition, the delay in revealing the outcome of the audience's action intensifies the effect that spectators seem to reflect only a little and then continue playing. In phase one, the relationship between the player's actions and the outcomes is discernable, but the long-term consequences of their actions are not clear. In phases two and three, the immediate results of the player's actions are not shown at all, and the long-term consequences are also not considered. As they do not know what the outcome of their actions will be, they cannot feel directly responsible for the results. The moments of reflection thus remain concise. When they do get confronted with the result of their choices, these results originate from actions earlier in the performance. The feedback loop which can create

moments of reflection is thus a long one. As a result, the audience may feel responsible for the way it acts but may not feel that it can change much because of the delay in outcomes. Here, it can be argued that this result is very similar to the operation of a real algorithm because the effect of an algorithm is mostly not directly visible, which makes it hard to change the effect.

Because the spectators do not act or reflect thoroughly out of responsibility, but likely feel responsible for their actions several times during the performance, the reflection piles up. This accumulation is manifested by the spectators' reactions at the end when there is no next task. Initially, people laugh but when the outcast audience members are asked to not only step out of the group but leave the building, the group becomes silent. The spectators look at each other and their body postures show that they do not know how to behave. The silent moment feels awkward and far too long, and after the performance is over, the audience leaves quickly. Looking at the notion of response-ability which I defined as the ability to respond through action or active reflection on the side of the spectator, here, the performance creates space for response-ability. The audience is invited to reflect on what they did or did not do in the performance, and on *how* it came about that they acted as they did. This invitation is not created by the performance in the form of an after-talk with the audience or something else where they can act, but by leaving them by themselves and creating space to think. Ultimately, in this last moment of confrontation, response-ability is shaped in a way that, looking at the audience's reaction, creates a big response due to the accumulated reflection.

Knowing that the performance was about making an algorithm, the link is made to the invisible mechanism and effects of an algorithm. Therefore, *#2 BIAS* provides insights related to the real world and intervenes at the level of perception by making the audience actively reflect on the theme of algorithms.

4. CONCLUSION

In this thesis I have investigated how shifting modes of audience address in *#2 BIAS* construct response-ability. The analysis shows that the notion of response-ability is constructed. The accumulation of moments of theatricality leads to response-ability. At the end of the performance, this accumulation results in an audience left with many things to think about and reflect on. By using the construction of an algorithm, the performance exposes the way an algorithm works, which most people are not aware of, and lets the spectators feel the consequences of that in a group setting. The dramaturgy of the performance copies the algorithmic principles, but simultaneously exposes them. The result is insight into the relationship between the operation of an algorithm and the negative consequences for others. Due to the shifting modes of address, space is created for response-ability.

What can we learn from this regarding algorithms? Where literature refers mostly to the online use and individual experience of algorithms, this performance provides a valuable live experience and a moment of reflection on the effects an algorithm can have on a group. However, the group dynamics and the ethical social dilemmas that arise within the group sometimes distract from the topic of the algorithm. Because your actions have a direct impact on others in the group, the focus can sometimes shift to how you should treat other people or how other people see you and vice versa. These 'distractions' could, however, also be seen as key issues within the topic of algorithms. *#2 BIAS* exposes the connection between the creator of the algorithm and the people who experience the outcome by giving the audience both roles. This highlights the transition from seemingly innocent opinions to outcomes that exclude people. The questions "how do you see other people?" and "how should you treat other people?" can criticise and improve this relation because these questions force us to uncover underlying biases as well as to ask which algorithm outcomes are desirable and which are not.

Another element that stood out in the analysis was the reason why spectators are excluded at the end of the performance. Literature shows that mostly marginalised people are excluded in different ways while in *#2 BIAS* people are sent away because they have 'leader qualities', or they are 'sexually attractive'.[51] These characteristics are usually perceived as 'good' in society and therefore mitigate the effect of exclusion.

This aspect is related to the goal of the algorithm, which is described in the performance as "making the group better". This goal is both abstract and seemingly meaningless: why should we make the group better? And 'better' by what standards? This vague goal again softens the impact of social exclusion. In real life, for instance, the goal of an algorithm is to find a candidate for a job; when

---

[51] Noble, *Algorithms of Oppression*; O'Neil, *Weapons of Math Destruction*.

someone is subsequently excluded from a job application by an algorithm, the setting is much more realistic and therefore painful.[52] This shows that how harmful the effects of an algorithm are depends on how much is at stake. Whether 'there is something at stake' is a subjective measurement in this matter.

The performance thus shapes the effect of exclusion but does so in a safely cautious way that makes the algorithm and its effects less realistic. Moreover, real algorithms are usually much more complex than *2 BIAS* implies, the performance thus simplifies real algorithms. However, the harmful effects of biased algorithms have been succinctly exposed and made tangible, precisely because it is simplified, and people understand the exclusion mechanism. The live aspect of theatre adds that a confrontation with these harmful effects cannot be avoided or escaped because you as an audience member are 'in' it.

2 BIAS, therefore, creates a critical perspective on the mechanism and effects of algorithms on a social group.

Follow-up research

Looking at this research, it would be intriguing to look at more performances that deal with algorithms, to compare how they address the audience, what problems they raise, and how much input the audience has. By comparing different performances, it would be possible to find strategies and insights regarding critical performances on algorithms. Interesting here, is how radical and real the theme of algorithms could be approached in a performance.

---

[52] Spice, "Questioning the Fairness of Targeting Ads Online."

BIBLIOGRAPHY

Bakshy, Eytan, Solomon Messing, and Lada A. Adamic. "Exposure to Ideologically Diverse News and Opinion on Facebook." *Science* 348, no. 6239 (June 5, 2015): 1130–32. https://doi.org/10.1126/science.aaa1160.

Blauw, Sanne. "Wat is een algoritme?" *De Correspondent*, July 2, 2019. https://decorrespondent.nl/10306/wat-is-een-algoritme/149980270484-745de161.

Bleeker, Maaike. *Thinking through Theatre and Performance*. Edited by Adrian Kear, Joe Kelleher, and Heike Roms. London; New York: Bloomsbury Publishing, 2019.

Bleeker, Maaike. *Visuality in the Theatre: The Locus of Looking*. London: Palgrave Macmillan UK, 2008. https://doi.org/10.1057/9780230583368.

Cormen, Thomas H., Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. 4th ed. Cambridge: MIT Press, 2022.

Flaxman, Seth, Sharad Goel, and Justin M. Rao. "Filter Bubbles, Echo Chambers, and Online News Consumption." *Public Opinion Quarterly* 80, no. S1 (2016): 298–320. https://doi.org/10.1093/poq/nfw006.

Fry, Hannah. *Algoritmes aan de macht: hoe blijf je menselijk in een geautomatiseerde wereld?* Translated by Johannes Jonkers. 2nd ed. Amsterdam: De Geus, 2018.

Groot Nibbelink, Liesbeth, and Sigrid Merx. "Dramaturgical Analysis: A Relational Approach." *FORUM+* 28, no. 3 (October 1, 2021): 4–16. https://doi.org/10.5117/FORUM2021.3.002.GROO.

Hiemstra, Annemarie, and Isabelle Nevels. "Algoritmes leiden niet automatisch tot eerlijkere selectie." *Sociale Vraagstukken*, December 3, 2018, 4.

Hoving, Roelf Anton. "Verdacht door een algoritme: Kan predictive policing leiden tot een redelijke verdenking" 7 (September 2019): 530–46.

Lehmann, Hans-Thies. *Postdramatic Theatre*. Translated by Karen Jürs-Munby. London: Routledge, 2006.

NEMO. *Bits of You*. 2021. Exhibition.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York university press, 2018.

O'Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin Books, 2016.

Pariser, Eli. *The Filter Bubble: What the Internet Is Hiding from You*. London: Viking, 2011. https://archive.org/details/filterbubblewhat0000pari_z3l4.

Phelan, Peggy. "Marina Abramovic: Witnessing Shadows." *Theatre Journal* 56, no. 4 (2004): 569–77. https://doi.org/10.1353/tj.2004.0178.

playField. *#2 BIAS*. 2021. Theatre.

Ridout, Nicholas Peter. *Theatre & Ethics*. Theatre &. London: Palgrave Macmillan, 2009.

Salen, Katie, and Eric Zimmerman. *Rules of Play: Game Design Fundamentals*. Cambridge, Mass: MIT Press, 2004.

Spice, Byron. "Questioning the Fairness of Targeting Ads Online." News Carnegie Mellon University, July 7, 2015. http://www.cmu.edu/news/stories/archives/2015/july/online-ads-research.html.

Sunstein, Cass R. *#Republic: Divided Democracy in the Age of Social Media*. Princeton: Princeton University Press, 2018.

Topol, Eric J., and Abraham Verghese. *Deep medicine: how artificial intelligence can make healthcare human again*. First edition. New York: Basic Books, 2019.

Vetzo, Max, and J. H. Gerards. "Algoritme-gedreven technologieën en grondrechten." *Computerrecht*, no. 1 (February 2019): 10–19.

Wright, Bradley R. Entner, and C. Wesley Younts. "Reconsidering the Relationship between Race and Crime: Positive and Negative Predictors of Crime among African American Youth." *Journal of Research in Crime and Delinquency* 46, no. 3 (August 1, 2009): 327–52. https://doi.org/10.1177/0022427809335170.