Title
Paramediation: Algorithmic Governance as the Parametric Design of Perceptible Phenomena

Author
Polack, Peter Joseph

Publication Date
2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Paramediation: Algorithmic Governance as the Parametric Design of Perceptible Phenomena

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Information Studies

by

Peter Joseph Polack

2022

ABSTRACT OF THE DISSERTATION

Paramediation: Algorithmic Governance as
the Parametric Design of Perceptible Phenomena

by

Peter Joseph Polack
Doctor of Philosophy in Information Studies
University of California, Los Angeles, 2022
Professor Leah A. Lievrouw, Chair

This study presents a theory of paramediation: the process by which algorithms govern human activity by arranging perceptible phenomena, according to parameters that structure how these phenomena emerge in correlation with human activity. It begins by identifying a tendency in existing research to conceptualize algorithms as representational and regulatory media, and it identifies how alternative theoretical conceptions of algorithms challenge this tendency by developing a phenomenology of algorithm operations. The theory of paramediation is proposed to analyze this phenomenological dimension, how it is designed, and its implications for the way we conceptualize algorithmic governance. Through illustrative case studies, participant interviews, and the production of an interactive game, the study interrogates implications of a theory of paramediation for studying, designing, and critiquing algorithmic media. It concludes that the power of paramediation cannot be checked by evaluating its epistemic validity, but by developing a plurality of practices for interpreting algorithm logic and data.

The dissertation of Peter Joseph Polack is approved.

Dana Cuff

Davide Panagia

Shawn G. VanCour

Leah A. Lievrouw, Committee Chair

University of California, Los Angeles

2022

Table of Contents

List of Figures and Tables

Vita

Peter Polack's research addresses the political implications of the theories, concepts, media, and social practices that we use to make sense of algorithmic media. He has published writing that forwards different theoretical and methodological approaches to studying algorithmic systems and their social consequences, and has exhibited artwork that narrativizes political conflicts engendered by these systems. During his time at UCLA, he participated in community education and organizing initiatives at the Stop LAPD Spying Coalition, where he collaboratively developed ideas and materials that informed his research. He holds an M.S. in Computer Science, a B.S. in Computational Media, and a Minor in Computing and Intelligence from the Georgia Institute of Technology.

Chapter 1: Introduction

Current efforts to regulate the social consequences of algorithms depend on identifying whether the design of algorithm logic accords with liberal values of equal representation (Ntoutsi et al., 2020), freedom of choice (Liberini et al., 2020), accountability (Diakopoulos, 2015, 2016), participation (Lee et al., 2019), and transparency (Diakopolous & Koliska, 2017). Historically, these principles have operated as guarantors of human liberties, provided a specific understanding of the relationship between individuals and the state, and they are readily applicable to understanding and regulating algorithmic decision-making processes. But these principles are also beholden to specific assumptions about how technical media operate in relation to human subjects. While heuristics of representation and transparency are suited to identifying whether media represent the world truthfully or accurately, they are inadequate to addressing modes of perception and interaction that algorithms facilitate. Altogether, despite its accessibility and popularity, the notion that algorithm code operates like a code of law (see for example, Lessig, 2000), and that its social consequences can be managed as such, is a deceptive one.

In this thesis, I argue that what is missing from this purview is an attention to how algorithms operate to arrange perceptible phenomena. Instead of taking for granted that algorithms govern human activity by regulating it – restricting or filtering what people can see, say, or do – we must attend to their capacity to arrange perceptible images, events and narratives to influence activity without restraining it outright, and namely to how this capacity is designed and implemented. This raises distinct issues for confronting the social consequences of algorithms, which are inapprehensible to current modes of criticism. It motivates us to identify alternative methods, registers of perceptibility, and forms of knowledge capable of apprehending its effects.

To account for this, I propose a theory of *paramediation*: the capacity of algorithms to arrange phenomena according to designed parameters, which specify how to sample information encoded in data, in order to elicit particular sensations and perceptions from people. The connotation of "para-" here is of something that happens alongside or parallel to something else, which remains correlated to it even as it appears to go beyond and alter it. Just so, algorithms possess the unique capacity to synthesize artificial percepts that correspond to material circumstances, and also to alter and optimize them according to specific parameters and heuristics, such as the extent to which they elicit certain human actions. What is unique about this process in comparison to other technical media is the specific way that phenomena are arranged: they are calibrated according to parameters, heuristics, or target functions that shape their emergence partially and indirectly. This exhibits a practice of *parametric design*, where constraints or parameters generate emergent phenomena that satisfy these constraints, rather than controlling them directly and totally. Such a practice is not concerned with the epistemic validity of the phenomena it produces, but with the pragmatic effectiveness of parameters and hyperparameters: how designed and inferred constraints guide the procedural emergence of phenomena, and how these phenomena correspond to human action in turn.

Examples of paramediation are numerous. They include the design of notifications to correspond to social media activity, the design of deepfake images to correspond to photographs, and the design of crime prediction maps to correspond to information about crimes. To call these notifications, deepfakes, and crime maps paramedia is to acknowledge that, while they are derived from data that are thought to have an unambiguous truth value, they produce something new from these data which introduce new, and yet controlled possibilities for meaning-making and action. What these notifications, deepfakes, and crime maps make sensible is, in a sense, otherworldly: it is designed to correlate to circumstances that are perceived as familiar or factual, but the perceptions they evoke cannot be said to have ever existed in the same way. They are articulated out of material traces that have been encoded in data, and then arranged according to specific parameters that optimize how they appear – characteristically to unanticipated effects.

Typically paramediation is carefully designed to realize particular tasks, like synthesizing images, directing attention, or producing coherent narratives, by arranging phenomena according to heterogeneous sources of data. This process is designed and evaluated according to heuristics that optimize its pragmatic effectiveness: namely, the extent to which emergent phenomena elicit certain actions in human subjects. Moreover, as in parametric design, because these emergent phenomena are irreducible or ancillary to the heuristics that shape them, unanticipated qualitative effects emerge. Consequently paramediation may manifest to people as *the algorithmic paranormal*, or uncanny synthetic figurations that betray their artificiality. Attending to these paranormal aesthetics, or paraesthetics, can help us to identify when paramediation is at play, and precisely how it is designed to organize appearances. Rather than identifying whether data or interfaces represent the world accurately (mimesis), or lauding their ability to augment human actions with tailored abstractions and automations (prosthesis), we examine how algorithms enable the arrangement of appearances according to data, as well as the optimization of these appearances in order to evoke particular sensations (paraesthesis).

Where algorithms are typically evaluated according to whether they correspond to a material or ideal truth (does this logic reflect reality? does this logic reflect the reality we want?), paramediation points to the capacity of algorithms to make something new appear out of the given. Neither definitively true nor false, the arrangements of appearances that result from paramediation, or *paramedia*, correspond at once to material traces encoded in data and to designed articulations of their details – a derivation and a deviation. To be sure, how material traces come to be accepted as given, as *data*, depends on each circumstance – data may be accepted as a given following a scientific method or a subjective observation – but the significance of paramediation is that it leverages this status of givenness to make something new appear, which inherits a part of what is given. From this view, the question is not whether algorithmic operations correspond to givens, but *how*: how phenomena are arranged in correspondence with the given, the familiar, the legitimated – in short, with *data* – to evoke particular perceptions and actions.

This understanding challenges a regime of algorithm criticism and reform that insists on interrogating the epistemic foundations and underlying ontologies of algorithm logic. Beyond an epistemological critique of algorithm logic, data, ontologies, and norms, we require a phenomenological analysis; and namely a post-phenomenological one that concerns how algorithms are *designed* to manifest to human perception. That is, rather than using phenomenology to study how subjects perceive algorithms (or how algorithms perceive subjects), we use it to understand how phenomena are synthesized and calibrated by design, namely by measuring, calculating, and optimizing their effects.

An attention to this phenomenological dimension of algorithm design motivates a shift in our understanding of what it means to be governed by algorithms: instead of viewing algorithms as procedures that aim to accurately represent or govern the world, we attend to how algorithms and human action are *designed together* by way of arranging appearances that correlate to human activity. Here, algorithm design is not just the specification of procedural rules to model and regulate subjects, but the strategic architecting of perceptible phenomena, where algorithm logic and human perception are both objects of design. This marks a shift from a concern that the ontologies underlying algorithm logic are partial or flawed, to a concern that algorithmic phenomena can be extraordinarily effective in their own right – and this is where algorithms derive their power.

Such a shift is not meant to disavow that algorithmic logic, computational ontologies, and data reflect partial design agendas and perspectives, but to insist that we do not approach equitable governance by making this logic more complete. Algorithmic governance is not just a matter of *inhibiting* human freedom with decisions informed by opaque logic and partial abstractions, but of *exhibiting* opportunities for meaning-making and action by calibrating appearances in relation to data. It is innovation in this area of algorithmic governance, in paramediation, that I argue current approaches to algorithm regulation and critique not only overlook, but characteristically incentivize.

Today, a focus on the decision-making logics and abstractions of algorithms cements a dependency on paradigms of governance, subjectivity, and media criticism that treat the impartiality, transparency, and flexibility of decisions as a limit to oppression. This motivates appeals to ensuring better representation of minoritized populations in data, transparent design of algorithm operations, and further end-user control over algorithmic decisions, which are approaches that have failed to address the social consequences of design principles underlying algorithmic systems (Polack, 2020). I argue further that these very approaches to reform depend on displacing the power of algorithms from regulation to paramediation, which raises new – and less salient – problems for social justice. These new problems are inapprehensible to traditional paradigms of algorithm criticism, insofar as they treat appearances of algorithms as evidence of underlying ontologies and decision-making logics, rather than strategically arranged phenomena in their own right.

This is reflected in, namely, recurrent proposals for new law enforcement information systems, which are explicitly branded to ameliorate social concerns raised about their predecessors. Branding in this case is not simply a matter of clever rhetoric or graphic design, but of designing algorithmic systems from the ground up to articulate effective paramedia: crime maps, automatically dispatched crime alerts, or even perceptible chains of causality between crime reports and police responses. To demand more complete data or less partial algorithm logic in such cases is to overlook how thoroughly these systems are designed to legitimate particular arrangements of power, authority, and social practices – irrespective of the robustness of data at their disposal. A theory of paramediation, in contrast, draws attention to how algorithmic configurations of perceptible phenomena operate to establish what the political philosopher Jacques Rancière calls a "distribution of the sensible," an arrangement of appearances that substantiates particular hierarchies, regimes of exclusion, and exploitative practices – making them appear taken for granted (Rancière, 2010). Notably, paramediation works less by excluding certain phenomena or by suppressing epistemic contradictions between them, than by incorporating them together in such a way that optimizes particular effects.

If paramediation raises any normative criterion for algorithm design or critique, then, it is to identify whether algorithmic media prioritize the *coherence* versus the *plurality* of practices for interpreting algorithm logic and data. While algorithms are particularly good at making coherent sense from data, by correlating these data, arranged phenomena, and human action, this may come at the expense of interpreting these data in different ways. In principle, algorithmic media lend themselves to operationalizing plural processes of making sense, because they can derive a wide range of possible phenomena, and thus possible interpretations, out of even the most unequivocal data. However, in practice they are typically implemented to streamline a particular, coherent and singular interpretation in order to substantiate a specific ordering of the world. Centering plurality against this tendency is not an appeal to collecting more data – which in fact tends to substantiate and reinforce singular sensemaking practices, by enlisting these data into its arrangements, however arbitrary they are. Nor is it a naive appeal to fragmenting sensemaking practices so thoroughly that they can never be reconciled. Rather, it is an acknowledgement of the way that a plurality of sensemaking processes is afforded by algorithmic operations, but also exploited by designers. If algorithmic systems are not designed from the ground up to admit this plurality, they deserve our utmost skepticism.

In this thesis I argue that certain practices already attend to these implications of paramediation, and that we can learn from these practices in order to perceive and evaluate the consequences of algorithms in new ways. In particular, I identify four settings where I hypothesize attention to paramediation is prominent: in grassroots community education initiatives about the social consequences of policing algorithms, in approaches to producing art about computational procedures, in the emerging field of user experience design for algorithms (algorithmic experience design; Alvarado & Waern, 2018), and in game design. I hypothesize these settings exhibit an attention to paramediation because they concern the conditions of algorithm appearance and their lived experiences for specific subjects – vulnerable populations facing criminalization, online platform users, art spectators, and gamers – who are (for very different reasons) the first to perceive consequences of algorithms in terms of what they make perceptible. This is to say that in community education, artmaking, user experience design for algorithms, and game design, organic accounts of paramediation and its consequences are being developed.

Confronting each of these practices and their considerations in turn, this study spans a wide breadth of disciplines. Nonetheless, its theoretical stakes are based in *information studies*, to which it contributes a theory of how information can be organized after its collection and formatting as data. This focus on the organization of information gives the study some methodological affinity with the field of *software studies*, which inquires into the design principles, materiality, and operations of specific computer systems. But most of where the study sources its background research and literature comes from a field called *critical algorithm studies*, which involves a variety of disciplinary approaches to addressing the social and technical nature of algorithms, as well as the discourses that attend them. In particular, I am specifically focused on the ways that this work is brought to bear on practical approaches to algorithm design and critique. The practical implications of this field are my theoretical point of departure: first identifying some of the main implications, then developing a theory that responds to a tension I identify in them, and ultimately using a variety of methods to evidence this theory. Overall, this theory is meant to first identify, then challenge a tendency in theoretical work that exerts an influence on practice, on the way to developing an alternative model for design and critique.

To accomplish this, I begin in Chapter 2 by surveying existing concerns about the social consequences of algorithms, as well as approaches to policymaking and design that address these concerns. Through this survey, I identify the theoretical premises that motivate these concerns and solutions, and problematize some of their assumptions. I show that while there is a growing attention to how algorithms are perceived by their subjects and audiences, these perceptions are regarded as a limit to algorithmic power, rather than an inherent aspect of it. In the second part of Chapter 2, I survey research and literature that concerns phenomenological analyses of algorithms, which sets the stage for developing a theory of paramediation.

In Chapter 3, I introduce the theory of paramediation. To describe the theory and identify questions for further investigation, I conduct an illustrative case studies of four algorithmic systems: generative adversarial networks used for image-to-image translation, the TikTok content feed, the crime data analysis platform Palantir Gotham, and procedural generation in the game Dwarf Fortress. The case studies are designed to illustrate what paramediation looks like across disparate uses of algorithms, as well as the aspects of algorithmic systems that a theory of paramedia enables us to attend to. I conclude this section by proposing some provisional principles involved in the design, realization, and appearance of paramedia.

The illustrative case studies serve as a point of departure for Chapter 4, where I describe an investigation of four corresponding settings of working with algorithms: artmaking, algorithmic experience design, community education about algorithms used in law enforcement, and game design. Through interviews with practitioners that work with algorithms in each of these fields, in Chapter 5 I investigate how the aspects of paramediation I identified in Chapter 3 accord with their work and understanding of algorithms. Further, I identify some of the unique ways that these practitioners attend to appearances and perceptions of algorithms. These interviews are designed to inform a theory of paramediation: they assess how the theory accounts for considerations involved in practical work with algorithms, as well as how these considerations contribute to a different understanding of algorithms.

Finally, in Chapter 6, the interviews inform the design of an interactive game that I produce to implement considerations of paramediation. I use the process of designing and implementing this game to examine the process of designing paramediation firsthand, as well as to assess the capacity of a custom-designed algorithmic system to demonstrate theoretical considerations and effects of paramediation. To evaluate the game, I conduct an additional round of interviews in which participants playtest the game and discuss its features from the perspective of their own disciplinary experience. In Chapter 7, I discuss the results of the analysis from Chapters 3 through 6, limitations of the study, and its theoretical implications.

Chapter 2: Literature Review

2.1 Trends in Algorithm Reform and Regulation

To define algorithm, we would do well to take a page from Ludwig Wittgenstein's playbook: a word's meaning depends on the specific ways that it is used (Wittgenstein, 2009); only through identifying how people use the word algorithm can we comprehensively grasp its meaning. Of course, we should also admit that there are things called algorithms at work irrespective of our awareness of them, underlying the computer systems that we interact with. This enables the philosopher of technology Yuk Hui to posit the existence of "digital objects" that owe their existence to their material configuration, and not to human creation and consciousness (Hui, 2016). But these two definitions of algorithm – the social and the technical – always appear to be at odds. They force us to acknowledge that the idea of the algorithm conflates two distinct, but interdependent referents: on the one hand, an algorithm is a digital object that operates independently of us, and on the other hand, an algorithm is a matter of concern, an object that exists insofar as we use or care about its effects (Latour, 2004).

This past decade, the corpus of research and literature called critical algorithm studies shifted the notion of the algorithm toward its social definition, namely by demonstrating the involvement of algorithms in a variety of existing matters of concern, from healthcare, to criminal justice, to electoral politics (Gillespie and Seaver, 2016). Evidencing this shift, Google searches of the word "algorithm" alone have declined since 2004, always rising and falling with the temporal cycles of academic scheduling, but searches for "Facebook algorithm" and "Instagram algorithm" have spiked dramatically in the last few years, despite the existence of these platforms – and their use of algorithms – many years prior. Unsurprisingly, the most significant spikes correspond to announcements from each of the platforms that they would be using new algorithms to determine the visibility of posts. Here, what appears to define the recent cycles of popular attention to algorithms is an interest in understanding the involvement of these algorithms in specific everyday activities that concern people.

Meanwhile, in the past decade scholars have started to use the term "algorithmic governance" in reference to the capacity of algorithmic technologies to automate or inform managerial decisions made on behalf of people. The term is used to stress the consequences that arise from granting algorithms the authority to make these decisions, which are said to mitigate or circumvent existing regimes of human authority and judgment (D'Agostino and Durante 2018). Algorithmic governance is therefore less of a concrete, localizable configuration of decision-making than a problematization of how algorithms supersede or compromise existing governance practices. Scholars of algorithmic governance stress that this supersession threatens institutional authority to make qualitative judgments (Crawford 2019), degrades transparency measures (Coglianese and Lehr 2019), and limits opportunities for people to engage in public decision-making practices (Danaher 2016).

In the past few years, journalists and news commentators have also started to use the term "algorithmic governance" to refer to dystopian and typically broad observations about the consequences of delegating decision-making processes to algorithms (Bridle, 2018; Mozur, 2018; Malik, 2019). It is in this context that, in April 2020, an entry for "algorithmic governance" was added to Wikipedia. In recounting the recent history of this term, I am not interested so much in demonstrating the emergence of a new regime of governance as I am in accounting for the popularization of a certain matter of concern. Indeed, such a concern with the "threat of algocracy" (Danaher 2016), along with the concerns of algorithmic bias, complexity, and opacity, may even draw attention away from how algorithms are actually designed to be embedded into the lives of people. While we should first question the narratives that inspire these trepidations (Ziewitz 2016), we should then investigate specifically how it is that algorithms are apprehended as matters of concern.

In this section I begin by describing four areas of concern about algorithms that have surged in popularity over the last decade: concerns that algorithms are biased, that they are opaque to human perception and consent, that they limit human participation in decision-making processes, and that they inhibit individual self-determination. Each area of concern motivates particular solutions to algorithm regulation.

2.1.1 Bias and Equality

In academic research, literature, and journalism, "algorithmic bias" refers to consequences of computational processes that are statistically partial to certain individuals, populations, or entities (Haijan et al, 2016). Such a bias can be attributed to either a technical cause, such as when an algorithm erroneously treats a certain data variable as more significant than another, or to a social cause, such as when an algorithm deployed in a particular social context exhibits an "emergent bias" that was not accounted for in the algorithm's design (Friedman & Nissenbaum, 1996). Algorithmic bias can also be attributed to the biases of input data (Fienberg, 2017), which is a relationship popularized by the phrase 'garbage in, garbage out.' For example, healthcare data that is disproportionately representative of a particular social group can cause health data analysis solutions to fail at operating for unrepresented groups (Garcia, 2016; Hague, 2019). This has motivated algorithm scholars to scrutinize the "preexisting" biases of individuals and social organizations that inform an algorithm's design and decisions (Friedman and Nissenbaum, 1996; Noble, 2018).

The prevailing response to algorithmic bias has been an appeal to detecting and correcting for algorithmic bias (Polack, 2020). This has motivated the development of "reverse engineering" methods (Diakopoulos, 2014), or statistical techniques designed to infer whether algorithms exhibit a bias toward certain entities. In almost all circumstances, reverse engineering is intractable due to the complexity of algorithms once they are operational in practice (Gillespie, 2014), and the epistemic claims of reverse engineering methods are generally dubious (Green et al., 2009).

But more significantly, appeals to detecting and correcting for algorithmic bias imply that bias is a property that algorithms either possess or do not. This overlooks that an algorithm always entails biases insofar as its decision-making criteria are not arbitrary – the only algorithm that does not exhibit any biases is one that is absolutely random, hardly an algorithm at all. Therefore, like the term algorithm, algorithmic bias tends to point less to a specific technical object than to a matter of concern: any discussion about algorithmic bias is in fact a matter of choosing which biases are concerning and must be controlled, and which to let pass. This explains why some appeals to evaluating algorithmic bias propose to compare an algorithm's decisions to those of humans, so as to permit algorithmic bias insofar as it is not more biased than a human benchmark (Cowgill & Tucker, 2017).

Altogether, algorithmic bias expresses the concern that algorithms threaten social standards of equality among people – either by compromising or circumventing human principles of equality, or by inheriting and exacerbating existing inequalities. These concerns are met with solutions that strive to ensure that the data and operations of algorithms regard people impartially and equitably. However, the algorithms that raise these concerns are most commonly those that are designed in the first instance to treat people differently on the basis of their personal characteristics and behaviors. In turn, this motivates criticism to forward a special formulation of equality where algorithms are permitted to treat people differently, but not on the basis of certain data points that raise concerns about discrimination.

In healthcare, for example, the correlation between socioeconomic status and health outcomes is well-documented (Adler et al., 1994). There is an ongoing debate about whether individuals with low socioeconomic status should be prioritized in receiving medical treatment to compensate for this fact; some scholars argue that such compensation mechanisms would conflict with the task to provide medical assistance on the basis of need and health risk (Hurst

2009; Bærøe & Bringedal 2011). Therefore, in order to avoid depending absolutely on socioeconomic status to allocate aid, research proposes to record, analyze, and monitor "community vital signs," or the social and environmental factors that put individuals at risk for illness (Bazemore et al., 2015). With recourse to "social determinants of health" (Marmot, 2005) that blur the social and the environmental, an analytical dependency on the category of socioeconomic status is avoided.

This formulation of equality rests on the strategic selection of variables that should influence an algorithmic analysis, in order to avoid an analysis of particular social categories. This is reflected more explicitly in reforms that the Los Angeles Police Department proposes to implement: a regime of data-driven policing that is not "discriminatory" or racist, but rather "discriminating"; that is, informed by data about crime and disorder in urban environments (Bratton 2018). The difference between discriminatory and discriminating policing is not just a rhetorical sleight of hand: it calls for algorithmic judgment based on behavioral and environmental factors as opposed to discrete social categories like race and class alone. Algorithmic systems will always depend on some criteria to differentiate among individuals; nonetheless there is a tendency to expand their scope of analysis in order to avoid explicit reference to discrete social categories. In the words of the policing scholar Andrew Ferguson, in a leaked email correspondence with the inventor of the PredPol predictive policing platform, "as a strategic matter you should promote your papers showing that you can balance race or other factors as a technical matter and it is all about how the police (not the companies) choose to calibrate the algorithm" (Request 20-5033).

The move to base data analysis on behavioral and environmental factors, at the ultimate discretion of analysts, as opposed to just social categories like race or class, does not prevent algorithms from treating their subjects inequitably. Indeed, it can inaugurate regimes of inequality that are less salient and more pervasive, drawing from sources of information that are more innocuous. This enables algorithms to appear as if they are less "discriminatory," a marker of inequality characterized by a clear relationship to social categories. By avoiding this categorization – by shifting from "discriminatory" to "discriminating" in the language of the LAPD – data collection pervades other aspects of life to derive indicators that cannot be so easily attributed to non-algorithmic forms of discrimination like human prejudice. That an algorithm treats social categories equally does not mean that it helps to ensure their equality, but it does help to obscure any inequalities that may persist among them.

Altogether, this drive for non-discriminatory data is supported by a liberalist conception of equality: individuals may be treated differently depending on miscellaneous factors, but they should be treated equally with respect to a governing logic that applies to them all. This formulation – an equality before algorithmic law – amounts to demanding that algorithms do not discriminate by salient social categories like race, gender, or socioeconomic status; instead algorithmic rules and logic should apply to all of these groups equally. Effectively, this is to view the standardization of algorithmic rules and conditions as more important to ensuring equality than the ultimate effects of these rules. By regarding algorithms as legislative rule sets, their improvement can be treated as a matter of designing the most equitable rules.

2.1.2 Transparency and Consent

Much like algorithmic bias, the idea of algorithmic transparency implies that if certain aspects of algorithms were made visible, then their problems could be more readily identified and corrected (Diakopolous & Koliska, 2017), especially by people who are most affected by them. Inversely, algorithmic opacity concerns how algorithms or their designers conceal algorithmic logic or its consequences from human apprehension (Pasquale, 2015; Burrell, 2016). The term implies a power asymmetry between people and algorithms: if the subjects of algorithms cannot witness how algorithms operate, then these individuals are less equipped to interpret and critique the consequences that algorithms have for them.

Algorithmic opacity expresses a concern that algorithms pose problems for hermeneutical justice. The philosopher Miranda Fricker proposes the term hermeneutical injustice to account for situations in which a person's experiences are rendered invisible or unrecognizable because the means, concepts, or words that could be used to make sense of the situation do not exist, or are not readily accessible (Fricker, 2007). In turn, Gwen Ottinger has demonstrated how hermeneutical injustice can be confronted by citizen science initiatives, where communities develop their own means of measuring and interpreting their experiences with industrial environmental hazards (Ottinger, 2017). The invisibility of these hazards and their denial by corporate stakeholders, despite their salient effects on victim's bodies, contribute to a hermeneutical injustice that renders victims less able to interpret the material causes of their own experiences. Just so, algorithmic opacity expresses the concern that we lack the resources, skills, and access to interpret how algorithms affect us.

In particular, algorithmic opacity problematizes algorithmic systems insofar as their consequences are impossible to anticipate or to attribute to particular design flaws. For this reason, existing approaches to algorithmic transparency focus mainly on making visible how an algorithm operates and what input data it ingests (Datta et al., 2016). As with algorithmic bias, algorithmic transparency implies that the discrepancies of algorithms might be mitigated by revealing their technical implementation details (e.g., which data variables it ingests, which machine learning methods it utilizes). However, since these algorithms are often proprietary and closed to public viewership, algorithmic transparency may seek compromises that disclose aspects of algorithms without exposing exactly how they function. In other cases, because operational algorithms are complex dynamic systems that are difficult to make sense of in practice, algorithmic transparency may work to make algorithms more legible to their developers, such as by developing visualization techniques or diagnostic tools that indicate an algorithm's state or decision-making processes (Sun et al., 2013; Bostock, 2014).

In any case, algorithmic transparency entails the fundamental assumption that opacity is an obstacle to understanding an algorithm's consequences. This assumption counterposes algorithmic transparency to algorithmic opacity, whereby transparency becomes a heuristic by which one can evaluate the epistemic or ethical validity of algorithms, irrespective of what exactly is being made transparent (see Ananny & Crawford, 2018). As with algorithmic bias, any appeal for algorithmic transparency against opacity is not comprehensive, because it is motivated by particular concerns about what should be accessible to human interpretation.

The opacity of algorithmic processes also raises concerns about their ability to make decisions on behalf of people without their awareness. This is reflected in research that theorizes how algorithmic control challenges existing paradigms of legal regulation and oversight. For instance, Shoshanna Zuboff's theory of surveillance capitalism problematizes the ways that algorithms break from the logic of the social contract (Zuboff, 2015). For Zuboff, the legibility of contracts affords laborers certain protections from capitalist exploitation, which are compromised by the automated rewards and punishments of algorithms that operate like "a new kind of invisible hand." She marks this shift as "the end of contracts," as algorithms replace "the rule of law and necessity of social trust as the basis for human communities." From this view, the visibility of contracts and mechanisms for consenting to them are safeguards compromised by algorithmic control.

A similar argument is taken up in Antoinette Rouvrouy's theory of algorithmic governmentality (Rouvroy & Berns, 2013). Rouvroy argues that the invisible influence of algorithms compromises the capacity of individuals to understand how they are being affected by them. This motivates her to insist on the importance of the law, and particularly on a legal conceptualization of subjectivity, to combat the invisible directions of algorithmic control. For Rouvroy, the language of the law enables a return to concrete, stable significations of selfhood and agency that algorithmic processing obscures.

At the basis of both of these theories is an insistence that algorithms should not act on people without their consent, or without some legal mechanism that can account for whether the relationship between algorithms and subjects is equitable. This premise begins from the assertion that algorithmic operations compromise the transparency and intelligibility of the law, which for its part establishes clear relationships of agreement and authority. Thus the premise rests on a distinction between algorithmic operations and legal protections, while at the same time holding them to the same standard of evaluation: insofar as algorithms do not operate with the visibility of the law – and according to the consent of the governed that this visibility enables – they compromise human liberties.

This mode of reasoning is prevalent in approaches to algorithm regulation. It includes campaigns to ensure that algorithmic services provide more intelligible processes of informed consent, like those notoriously obscured by Terms of Service Agreements. More significant are appeals to develop what Iyad Rahwan calls "algorithmic social contracts," or systems for monitoring algorithm operations according to their alignment with the values of stakeholders (Rahwan, 2018). An algorithmic social contract is a speculative interactive system designed to support algorithm regulation by bringing algorithm operations closer to the ideals of social contract theory, ensuring that algorithm operations correspond with what certain people expect from them. A key component of this approach is the concept of "society-in-the-loop," which is supposed to alter the operations of algorithms according to real-time feedback from its stakeholders. Here we observe a shift to interactivity in the name of preserving the logic of the social contract in algorithm design: by making algorithm regulation more dynamic and interactive, consent can be facilitated in real time.

These approaches to algorithm regulation reflect a growing tendency to acknowledge how people perceive (or are unable to perceive) algorithms. A complementary idea is that, by making the operations of algorithms more perceptible and interactive, people are less susceptible to being exploited by their rules. This view has been popularized by the strand of algorithm regulation called "trustworthy AI" (Floridi, 2019) which integrates the considerations of algorithmic bias, transparency, and accountability with an acknowledgement that algorithms can breach the trust, and the consent, of their subjects. In turn, the notion of trust in AI motivates the idea that algorithm operations should be designed to maintain the trust of their subjects. It thus shifts the terms of algorithm ethics from mitigating social harm and ensuring beneficience to ensuring consent.

Through terms of service agreements with broad language (Pollach, 2007), appeals to initial or tacit consent (Cohen et al., 2014), and an insistence on social contract theory (Zuboff, 2015), approaches to algorithmic transparency tend to exhibit a resoundingly legalistic conceptualization of algorithms, where a liberalist paradigm of consent determines the nature of transparency. As opposed to striving for the transparency of all aspects of algorithm logic, consequences, possibilities, or harms, transparency is sufficient so long as algorithmic operations are consented to. This logic is made explicit in appeals to "libertarian paternalism" (Sunstein & Thaler, 2003), which proposes to limit user control over algorithmic operations to those operations that concern them personally, unburdening them from the need to make extraneous decisions that can be delegated to algorithmic processes. So long as the system is designed to operate in the user's best interest, the consent of the algorithmic governed can be presupposed.

While an extreme case, the "freedom of choice" that libertarian paternalism aspires to afford users – precisely by limiting the choices available to them – once again demonstrates the fundamental tension of algorithmic transparency: what exactly should be made transparent? For libertarian paternalism, the answer has less to do with indicating the potential consequences of algorithmic operations, than with indicating operations that users can choose to consent to. Namely, this entails representing algorithms as transparent rule sets that users can modify, such that whatever happens beyond the purview of this "choice architecture" is already implicitly consented to, even if users are unaware of it. Such a formulation of transparency minimizes concerns about the configuration of algorithms and social relations that do not directly relate to conscious, individualistic concerns (Sharon, 2016). It favors a selective enclosure of algorithmic decision-making processes, supported by a logic of partial transparency and consent to opacity.

2.1.3 Accountability and Representation

Faced with concerns of algorithmic bias and opacity, "algorithmic accountability" (Diakopoulos, 2015, 2016) proposes to find ways to hold algorithm developers accountable for the consequences of algorithms. This includes the development of tools for detecting algorithmic bias or ensuring algorithmic transparency, so as to enable "algorithmic auditing" (Mittelstadt, 2016; Mehrotra et al., 2017). Generally appeals to algorithmic accountability propose to supplement the decision-making criteria of algorithms with human oversight (Neyland, 2016; Shneiderman, 2016), as in a system of checks and balances. For example, a risk score assigned to a criminal suspect by an algorithm might be evaluated or challenged by human judges. In this way, in contrast to algorithmic bias and algorithmic opacity, algorithmic accountability tends to focus on ensuring that people can evaluate and control the consequences of algorithms, as opposed to avoiding them in advance by changing an algorithm's design or input data.

Insofar as algorithmic accountability follows from the premises of algorithmic bias and algorithmic opacity, it inherits their presuppositions. For algorithmic accountability the aim is to control the consequences of algorithms by evaluating algorithms to ensure that they are epistemically valid. This commonly amounts to evaluating algorithmic mistakes; for example, instances where an algorithm yields a false positive or false negative by classifying something as what it is not. But because this scope of concern about epistemic validity takes issue with classification errors, it marginalizes the implications of classification itself; namely, that classification already exhibits partial schemes of ordering the world before this scheme is implemented in algorithms (Mol, 1999; Bowker & Star, 1999). This focus consequently prioritizes controlling algorithms by responding to their epistemic mistakes (i.e., false positive and negatives) as opposed to ensuring that epistemically valid operations (i.e., the detection of true positives and negatives) do not lead to harmful consequences themselves.

Altogether, the scope of algorithmic accountability is particularly limited to responding to the consequences of algorithms as they arise, as opposed to designing algorithmic systems with a sensitivity to their possible social impacts. Appeals to algorithmic accountability tend to follow from algorithmic mistakes, mishaps, or blunders that draw the reliability or epistemic legitimacy of algorithms into question once they are already deployed in practice, manifesting them as matters of concern. Algorithmic accountability thus intends to account for algorithmic discrepancies when they emerge in particular circumstances, but it does not venture to attribute these discrepancies to concrete design principles and assumptions that might reproduce them in the future. Therefore algorithmic accountability wants to hold algorithm developers accountable for the systems that they have constructed, but less so for the systems that they are actively developing.

Moreover, the ability to provision human oversight to detect these consequences is effectively reserved for algorithm developers, not algorithm subjects. While it has been argued that this unfairly places the onus of algorithm consequences on algorithm developers (Mittelstadt et al., 2016), it also makes the subjects of algorithms responsible for challenging the consequences of algorithms as they confront them and demanding more rigorous accountability measures, which are demands that will not necessarily be welcomed by algorithm developers (Kemper & Kolkman, 2019). Thus the notion of algorithmic accountability involves a dichotomy between algorithmic developers and algorithmic subjects, whereby the former are supposed to be held accountable by the latter's demands. While the notion of algorithmic accountability fundamentally addresses the concern that human participation in the decision-making processes of algorithms is limited by their opacity and automaticity, it also raises concerns about how inclusion and representation in algorithmic decision-making process is itself stratified.

But in practice, algorithmic accountability rarely amounts to extending participation in algorithmic decision-making. Instead, another type of inclusion is dominant, which follows a liberalist paradigm of representation where formal inclusion in a governance system amounts to full-fledged participation. We see this in the notion that more diverse populations should be represented in data and algorithmic procedures. For example, the media scholar Benjamin Bratton (2022) proposes to make algorithmic models more equitable via "quantitative inclusion," or including more people in the data they ingest. Meanwhile, Cohen et al. (2014) propose the use of graduated licensing fees to make health analytics services affordable for institutions with fewer resources. The task of algorithmic accountability to extend human participation in the design and evaluation of algorithms is subordinated to formalizing representation in data or algorithmic procedures.

The type of representation found in "quantitative inclusion" can be leveraged by campaigns that aim to enroll people into algorithmic services, like enrolling houseless people into a blockchain registries under the pretense of granting them access to healthcare and welfare services, while also maintaining a record of the service they receive (Hamburg & Collins 2010). Here the right to healthcare and welfare is presented as an incentive for opting into a data analysis system. This incentive is coercive insofar as an unwillingness to enroll in a welfare platform and cooperate with its requirements justifies the withholding of aid. Whereas a liberalist premise of inclusivity holds that no group of people should be excluded from algorithmic services, the corollary of this principle is that it justifies the neglect, and sometimes even punishment, of individuals who do not consent to them.

This formulation of inclusivity resembles conditional welfare policy, which aims to provide individuals with social welfare benefits so long as they abide to certain regulations (Dwyer, 1998). Conditional welfare personalizes benefits to individual behavior, which effectively personalizes eligibility for welfare. In the case of algorithms, eligibility may be calculated according to factors over which individuals have little control (Eubanks, 2018). Here the assumption is that giving individuals equal access to services makes them more equitable. But it effectively places the onus of proper behavior on the actions of individuals. Those who do not adhere to the conditions laid out by the welfare policy are excluded from enjoying its benefits, but paradoxically their virtual representation by the policy is presupposed.

Altogether, these appeals to inclusion in data and algorithmic inference exhibit a failure to think beyond a liberalist paradigm of representation. Much like removing social categories from analysis to avoid accusations of discrimination, this form of representation aims to use social categories as heuristics for fair treatment, without regard for algorithmic discrimination that happens by way of inference or conditional logic. Demands to represent individuals and demographics more equitably in algorithmic decision-making does not guarantee their participation in making these decisions, and may even justify their subjection to a conditional logic of compliance.

2.1.4 Influence and Responsibility

Another kind of apprehension about algorithmic governance concerns not whether algorithmic calculations are biased, opaque, or regulated by human oversight, but their very capacity to influence human activity. Such concerns address the role of "machine learning" and "artificial intelligence" in operationalizing an epistemology of statistical inference that treats statistical relationships among data as causal relationships in the world (for example Rouvroy, 2013). While the statistical methods underlying these techniques are not new, what is unprecedented about algorithms is their capacity to automatically shape the material and symbolic world in the image of their classifications, predictions, and ontologies.

This is demonstrated thoroughly by research on the "quantified self" (for example Swan 2009; Raghupathi & Raghupathi 2013), which characterizes self-tracking applications that collect and analyze data about people to display information about their health or physical fitness to them. Research on the quantified self addresses how algorithms present calculations to people that influence how they perceive themselves, and accordingly how they behave (Moore & Robinson, 2016; Sharon, 2017). Likewise, the algorithmic technologies underlying search engines, recommender systems, and targeted advertisements have come under focused scrutiny for influencing how people perceive relationships between ideas (Ananny, 2011; Noble, 2018), or for influencing human behavior directly (Yeung, 2017; Liberini et al., 2020; Ribeiro et al., 2020). In each of these cases, an algorithmic influence on human behavior raises concerns about human agency, or the capacity of individuals to act independently of external influence.

A common response to these concerns is to provide individuals with further control over their data. This is reflected in the EU General Data Privacy Regulation (GDPR[1]), which provides individuals with legal protections that enable them to request information about how their data is being analyzed, and also to request restrictions on this analysis. Crucially, however, while the GDPR grants EU citizens legal procedures to access and control data analysis, it also exempts data analysis from these regulations if it pertains to certain use-cases, like when it used "to ensure the quality and cost-effectiveness of the procedures used for settling claims for benefits and services in the health insurance system" (GDPR recital 42), or more broadly, when it involves "high public interest" (GDPR article 81(2a)). Despite its proposal to regulate algorithmic control, the GDPR prioritizes a broad formulation of public interest.

Consequently, the subjects of the GDPR legislation become responsible for requesting restrictions on these services, to the limited extent that they are authorized to. This enables data analysis services to use personal data liberally, while placing the onus of restricting this use on the persons themselves. What is more is that these people may depend on these services to receive access to resources and aid, and so may not have the luxury to opt out of them. This logic

---

[1] Any mention of the GDPR refers to Regulation 2016/679 of the European Parliament and of the Council on the Protection of Natural Persons with Regard to the Processing of Personal Data and on the Free Advancement of Such Data, and repealing Directive 95/46/EC (2016).

reinforces an individualist conception of subjects as rational agents who have the freedom to limit algorithmic power at their will – insofar as data analysis for public interest remains intact. It eschews considerations of dependencies that people have on algorithmic systems, and also of configurations of power and responsibility that limit individual choice.

This logic is also reflected in proposals for data propertization, which propose legal frameworks for treating personal data as private property, such that data cannot be taken from individuals without their intentionally giving it away or selling it. While such reforms enhance an individual's control over data collected about them, it also motivates them to treat the confidentiality of information about them as an asset, which can be sold. Logically, insurance companies already grant insurance premium benefits to individuals that consent to having their data collecting and analyzed (Barlyn, 2018). Here financial remuneration is not even necessary to operationalize personal data as a private asset that can be exchanged for better treatment by healthcare services. Data propertization maintains that people are free to keep or sell their data, while neglecting the distribution of power that constrains or coerces this ostensibly free choice.

These approaches to algorithm regulation suggest that, although algorithms make decisions on behalf of people, people are still in control – they retain the freedom to restrict automated decisions made on behalf of them. Here a liberalist notion of rational choice and property ownership takes precedence to an understanding of the agency delegated to automated algorithmic processes and their capacity to influence human decisions by design. To be sure, individuals are by no means in control of how their data is used and how algorithmic services act towards them, even when they have access to this information or the ability to cancel the process. Legal regulations like the GDPR and data propertization ensure that people take ownership of their profiles and the data analyses that produce them, even when these profiles and data analyses are devised and controlled by other parties. They effectively make individuals further responsible for challenging and regulating algorithmic operations, within established limits and bounds.

2.2 Precedents in Algorithm Phenomenology

Throughout research and literature, algorithms are seen to coerce human behavior by enforcing computational rules and norms that determine how people can act, think, and move. They are seen to "regulate" activity, either by reacting to it in real time or by predicting its trends to shape it in advance (Yeung, 2018). The most popular scholarship on algorithmic harms, biases (Eubanks, 2018), and governance (Rouvrouy, 2013) concerns these regulative coercions: norms or rules that inhere in algorithm logic and subject people to them directly, restricting their self-determination.

An *epistemological approach* to analyzing algorithms is applicable here. It concerns algorithms as *representational* or symbolic systems that represent entities and their relationships in the real world, and it concerns whether or not these representations are accurate and equitable. That is, if algorithm logic presupposes, represents, or implies the existence of certain entities and relationships in the world, an epistemic analysis can determine whether these entities and their associations exist in fact, or ought to exist. An epistemological view of algorithms holds that their logic either *represents*, *misrepresents*, or *reflects the existence of* real entities, and that their operations are valid insofar as they represent these entities as accurately as possible.

An epistemological analysis of algorithms can serve as a diagnostic for assessing whether algorithm logic is designed according to liberalist principles of governance identified in the previous section, such as equality under the algorithmic law, equal representation and quantitative inclusion in data, transparency, and tacit consent. These principles rest on a conceptualization of governance as that which must (1) represent the interests of individuals accurately, (2) comprehensively such that these interests are not violated, and (3) in a way that is transparent. Governance by algorithms is typically conceptualized in this way: algorithms represent individuals as data, and act on them in a way that could conflict with their interests or circumvent their awareness. From this view, algorithms operate mimetically and normatively, representing an objective reality according to norms that encode and organize data, and then using these data to regulate human activity. Thus algorithms are seen to constrain activity in the image of their logic, ontologies, and norms. This formulation implies that we can trace the consequences of algorithmic interventions directly back to their logic and data.

However, other formulations of algorithmic governance are possible. While concerns about discrimination, inclusivity, transparency, and agency dominate practical efforts to algorithm reform, algorithms are conceptualized in many other ways in theoretical research and literature: as media, as anthropological artifacts that produce meaning, and as agents of subjectivation, which raise different challenges for algorithm regulation and social justice. Altogether, this work suggests that the capabilities of algorithms are not limited to operationalizing ontological assertions about what exists or ought to exist – they also produce phenomena that become objects of analysis in their own right, and which subjects of algorithmic systems perceive. Such a *phenomenological* view of algorithms challenges the idea that the social consequences of algorithms are exhaustively reflected in their logic. By attending to this work, we can acknowledge some limits of existing approaches to social justice in algorithm regulation, and shift our understanding of what it means to be governed by algorithms.

2.2.1 Defining Ontology

An epistemological approach to algorithm criticism inherits some principles from computer science and information science; namely, a conceptualization of data and its formatting as a representation of particular subjects. Somewhat confusingly, this conception of data and its formatting as a representation is called an "ontology" in computer science and information science. Here the use of the term ontology has less to do with a philosophical interrogation of being than with establishing a correlation between computation and the subjects of data. It refers namely to a formalized symbolic system that can be encoded into computer logic. This means that computational ontologies are an aspect of an epistemological approach to algorithm criticism: identifying whether computational representations correlate to the subjects they aim to represent. This link between computational ontology and an epistemological approach to algorithm analysis should not be interpreted as implying that the epistemological approach involves an ontological inquiry broadly conceived.

We can distinguish between two conceptualizations of algorithm ontology, which give rise to different approaches to algorithm criticism and reform. A *hard* approach to ontology in computer science concerns the capacity of computer systems to identify actually existing entities in the real world, which effectively presupposes a universal reality independent from human observation, and a stable metaphysical ontology. A *soft* approach, on the other hand, is not meant so much to identify what exists in the world, but to pragmatically develop particular ontologies that congrue with, or approximate circumstances in the real world. The quality of these pragmatic approximations is informed by heuristics, which can measure the capacity of a computer system to interact on real-world circumstances effectively and pragmatically. (This account of ontology could be seen to resemble the philosophy of Rudolf Carnap, for whom ontologies are relative to specific symbolic systems or languages that establish them.)

Existing approaches to algorithm critique concerned with epistemology inherit either the hard or soft approach to computational ontology from computer science and information science. An approach to algorithm critique informed by hard ontology seeks to disprove that a given algorithmic system adequately represents actually existing entities in the objectively real world. An approach to algorithm critique informed by soft ontology seeks to demonstrate that an incongruity between an algorithm's ontology and the subjects of data lead to particular consequences. In either case, both approaches are concerned with the correlation between computationally instantiated ontologies and real-world circumstances; that is, whether a computational *representation* is adequate, according to certain criteria, and transparent in doing so.

In order to assess this, the primary object of analysis is the data formatted by algorithm code and decision-making logics, and how the content and form of this data correlates to an objective reality. A hard approach may tend to conceptualize data as an objective representation of reality, whereas a soft approach may hold a more pragmatic, albeit positivist view, where data is thought to represent an objective reality adequately or not. Through the triangulation of algorithm logic, data formatting, and real-world circumstances, either approach can demonstrate how algorithm logic fails to adequately represent an objective reality. This conceptualization is what motivates us to distinguish between an epistemological and a phenomenological approach to algorithm analysis: the former is concerned with the design of logic and the evaluation of data, as opposed to the capacity of a phenomenological inquiry to address other dimensions of algorithm design, use, and effects.

2.2.2 Interface Studies

The most common approach to what might be called an algorithm phenomenology rests on an analysis of how algorithms appear to human perception through *interfaces*. Usability studies in user interface design (UI) and user experience design (UX) approach already resemble an algorithm phenomenology, insofar as both are concerned with examining how people use, perceive, and understand computational systems (for example, Nielsen, 1994). These studies conform to a classical understanding of the interface as a standardizing protocol that makes individual units of code interoperable with one another. For them, the interface is what discloses certain computational functions to a user, or otherwise encloses them so that they cannot be accessed nor used.

By this account, the interface is conceptualized as a kind of gateway that selectively makes certain computational processes usable and visible, or otherwise makes them inaccessible and invisible (note that usability and visibility are virtually synonymous here). Thus the interface serves a mimetic function: it is supposed to accurately represent objects in the world or the computational processes that represent them in turn. Such theories of the interface as a representation also imply that it is a limit to perfect mimesis. The interface distorts the reality it feigns to represent, it reifies algorithmic logics, it is opaque. This is a typical source of consternation for research in gamification studies, which investigates how the procedures of information systems, algorithms, and data collection procedures are "gamified," or designed as entertaining games (Woodcock & Johnson, 2018). The purpose of gamification is to make ordinary or displeasurable interactions with information systems more entertaining, which can be read as a kind of deception at the level of the interface. For example, gamification research addresses how people are not remunerated for their labor because their labor is designed to appear as a game that is entertaining to them – what the videogame researcher Julian Kücklich calls "playbor" (Kücklich, 2005) and what the legal scholar Julie Cohen calls "playing and being played" (Cohen, 2014). Such research highlights the discrepancy between the consequences of algorithmic operations and their deceptive appearance to users.

Where usability studies aim to improve the effectiveness of interfaces to make human-computer interactions more efficient, a phenomenological approach would concern how people are able to understand computational processes despite their concealment. Seminal research in this area analyzes how people develop their own "folk theories" about how algorithms underlying their social media news feeds work, namely on Facebook (Rader & Gray, 2015; Eslami et al., 2015). Other work explores how subjects of algorithmic systems develop and operationalize these theories, such as gamers that develop theories about how games work through reverse engineering or "theorycrafting" (Wenz, 2013), and social media users whose online activity reflects a tacit knowledge of algorithmic processes (van der Nagel, 2018). A different line of research investigates human perceptions of algorithmic bias and fairness in everyday contexts (Lee, 2018; Woodruff et al., 2018). These studies challenge concerns about algorithmic opacity by demonstrating that, while algorithms are invisible to their users in one sense, users comprehend how algorithms work through their interactions with them.

An alternative to analyzing what interfaces disclose and make visible to people (representational analysis), or analyzing how people interact with and perceive algorithms through interfaces (phenomenological analysis), is to analyze interface design conventions, aesthetics, and systems of communication that are generated by algorithms (semiotic analysis). In a semiotic analysis, the interface is regarded as an expressive medium, like a film or a text, following from philosophical literature in media studies and textual analysis.

Although Lev Manovich is often credited with identifying the interface as a "cinematographic" medium characterized by its recombinations of spliced data (Manovich, 2001), we find early precedents of this account in Félix Guattari's theories of computational semiotics (Guattari, 1996), as well as in Vilém Flusser's distinction between the linearity of textual media and the mechanical procedures that composite "technical images" (Flusser, 2011). A cinematographic account of the interface is taken up once again by Maurizio Lazzarato (2014), who follows Guattari to propose semiotic and aesthetic relationships between cinema and interfaces broadly conceived. But these cinematographic accounts of the interface are relatively abstract, mainly concerning semiotic theories as opposed to specific interface design conventions; it remains up to later scholars to concretely analyze interfaces in these terms.

In this vein, Ganaele Langlois uses Guattari's semiotic framework to analyze how algorithms underlying Amazon's web store interface articulate semiotics that are designed to be perceived to people, like personal pronouns and book covers, to convey meanings to users (Langlois, 2008). Similarly, drawing from theories of materiality and performativity, Johanna Drucker proposes to investigate the "enunciative dimensions" of interfaces that appear to speak to users, and how they are articulated by algorithmic processes (Drucker, 2013). Drucker's theory of performative materiality demonstrates that the interface, like a written text, is performed by the interactions and interpretations of its users at the same time as its materiality constrains the probable range of these activities. For Drucker, interfaces are designed to solicit certain actions, enunciations, and interpretations from users, but they do not constrain this activity absolutely. This emphasis on interpretive flexibility and performativity challenges claims that interfaces reify algorithmic logic in the world, while it also leaves room for acknowledging the material constraints that interfaces impose on human conduct.

2.2.3 Algorithmic Semiotics

To account for the existence of algorithm and interface semiotic conventions that do not operate textually or cinematically, I proposed a theory of algorithmic semiotics (Polack, 2020). Algorithmic semiotics convey ideas, affects, and causal relationships to people by coordinating algorithmic events in space and time. This includes algorithmic "nudges" (Yeung, 2017) and "dark patterns" (Brignull et al., 2015) that respond to data collected about user activity with push notifications and subtle changes to interface layouts. Algorithmic semiotics can also operate like "procedural rhetoric," which refers mainly to the capacity of interactive videogame narratives to communicate ideas by dynamically responding to a person's choices (Bogost, 2007). A theory of algorithmic semiotics also encourages us to attend to the design conventions that configure distributed networked devices as interfaces, as in the "Internet of Things," or as what Christian Andersen and Søren Pold call the ambient and ubiquitous "metainterface" (Anderson & Pold, 2018). Algorithmic semiotics refer to the conventions, whether intentionally designed or incidentally generated, by which algorithms become capable of signification.

The notion of algorithmic semiotics supports a phenomenology of algorithms that exceeds the scope of interface studies: interfaces are not the only way that algorithmic logic becomes apprehensible and meaningful to people. One example of this is illustrated by Shintaro Miyazaki's theory of algorhythmics, or the temporal patterns that emerge from algorithm operations and interactions, which manifest in the material world and can be measured (Miyazaki, 2012). Somewhat like Bogost's theory of procedural rhetoric, Miyazaki sets his sights (or his hearing) on what the procedures of algorithms make apprehensible over time, through sequences of outputs or rhythmical patterns. For Miyazaki, these emergent phenomena demonstrate that the impact of computation on culture is inherently time-based, from the micro-level of processing bandwidth to the macro effects of algorithmic financial trading. The theory of algorhythmics encourages us to broaden the epistemic scope of algorithm studies, to acknowledge how and when their effects are apprehensible as temporal patterns and rhythms, and how the salience of these percepts informs how algorithms are perceived and analyzed.

2.2.4 Algorithmic Experience

Beyond investigating user "folk theories" or "theorycrafting" about algorithms, research also examines the less defined, tacit perceptions of algorithms that users depend on to make sense of their interactions with them. Taina Bucher's (2017) theory of "algorithmic imaginaries" points to the affective dispositions of users that emerge from their interactions with the Facebook algorithm, which condition their usage of the Facebook platform, and in turn inform the behavior of the Facebook algorithm itself. This includes user feelings that content recommended to them by algorithms is so accurate that it feels "creepy," a feeling examined in other scholarship as well (Shklovski et al., 2014; Phelan et al., 2016). The field of computing design and criticism called "postcolonial computing" emphasizes a similar relationship: everyday experiences with a technology condition its use, at the same time as the technology conditions these experiences (Irani et al., 2010). This "generative" model of computing use attends to the way that cultural differences and power dynamics can shape technology use and value irrespective of a designer's intentions.

Such a concern with the interactions between technology design, experience, and use has been examined extensively in science and technology studies (STS). STS challenges the idea that the design of technology and scientific inquiry evolves linearly to solve particular problems, without interference by social or cultural factors. Scholars in this field often examine how perceptions of technology influence their use and motivate changes to their design, as well as how the design of technology informs these very perceptions. For actor-network theory (Latour, 2007), for example, a given technological artifact is perceived as such insofar as it is "black boxed," or perceived in a way that simplifies its underlying complexity (see also Latour, 1987). Similarly, for social construction of technology (SCOT) theory, user experiences with technology influence its function and motivate changes to its design, until compromises between various stakeholders harmonize and arrive at "closure" (Pinch & Bijker, 1984). For SCOT, subjective interpretations of technology are heterogeneous and never totally determined by design, giving rise to an "interpretive flexibility" that influences an interplay between experiences and uses of technology (Pinch & Bijker, 1984).

Attending to experiences of algorithms in particular, Ned Rossiter and Soenke Zehle theorize an "aesthetics of algorithmic experience" (Rossiter & Zehle, 2015). Against the tendency of software studies to focus on the agency of computer systems and its determination of human activity, Rossiter and Zehle are concerned with human experiences of computation that escape these determinations – even as computational systems may be designed precisely to quantify and "capture" human experience. Rossiter and Zehle also critique the purview of actor-network theory and object-oriented ontology, which, following the media theorist Alexander Galloway's account (Galloway, 2018), reduce experiential and antagonistic encounters among subjects and objects to a flat network of connections that can be operationalized in the service of information extraction. Put another way, actor-network theory and object-oriented ontology use a conception of ontology from computer science to conceptualize all things.

A theory that echoes the notion of algorithmic experience is "algorithmic superstructures" (Anikina, 2020) which refers to the capacity of algorithms to intervene in the construction of human meaning, as opposed to intervening directly in the material reality which conditions these meanings. In her introduction to the art exhibit that introduced this term, the film artist and scholar Alex Anikina draws from Langlois and others to call "algorithmic superstructuring" the practice of designing algorithms to configure meaning in various ways, and calls on artists to reveal and experiment with these configurations. While the definition of material practices and techniques that enable algorithmic superstructuring is left open-ended, the term acknowledges the possibility of thinking algorithms beyond mere procedures, as designed interventions into the articulation of meaning and experience.

A different account of algorithmic experience called "AX" — a play on the field of user interface design called UX or user experience design — aims to account for how users experience algorithms, namely to design algorithms such that their experience is more positive and conducive to human needs (Alvarado & Waern, 2018). AX is an attempt to systematically account for user perceptions of algorithms in a way that goes beyond typical heuristics of usability from user interface (UCI) and human-computer interaction (HCI) design. Insofar as AX is concerned with user perceptions of algorithms, it marks a practical shift from an epistemological analysis of algorithm logic to a phenomenological analysis of algorithms.

2.2.5 Algorithms in Capitalism and Governmentality

Theories of algorithmic phenomena are intimately related to theories of algorithmic power; indeed, what people are able to perceive about and through algorithms informs their capacity to act in the world and exert an influence on others. Theoretical work that explores this relationship tends to follow from the philosophical frameworks of Karl Marx, Michel Foucault, Félix Guattari, and Gilles Deleuze, who are all concerned with the capacity of people to identify and react to phenomena that are arranged, either by social practices or technologies, to inform their behavior. The differences between these frameworks, and namely between Marxian and Foucauldian conceptualizations of technology design, reveal some of the concentrations of each.

Marxian theories, like Kücklich's theory of playbor (Kücklich, 2005), are predominantly concerned with the capacity of algorithmic systems to obfuscate the material interests and practices of their designers. These Marxian approaches to technical media criticism rest on a conceptualization of algorithms as automatic and mimetic media: algorithms enact a process that produces some value (usually economic), but the production of this value is obscured by a system that misrepresents it, thereby deceiving its users. It may be that gamification obscures the value production at work behind gameplay (Kücklich, 2005), that online platforms deceive us into thinking that their services are free while extracting our data (Fuchs, 2014), that partial data deceives us into thinking that its inferences are comprehensive (Joque, 2022), or that computation obscures differences between qualities by treating them as commensurable (Beller, 2021).

The emphasis on misrepresentation is inherited from Marx's method of analysis, which addresses how particular social practices and values come to be taken for granted in a way that obscures their historically contingent nature and their underlying inequalities (Marx, 2004). One reading of Marx insists that, beneath the appearance of these social practices and values, there is a material reality that should be uncovered to reveal the truth of mystifying and deceptive appearances. For scholars of algorithms and Marx, this formulation has the affordance of mapping on neatly to the notion of 'black boxed' information systems: the interface is a reified, deceptive appearance that obscures an underlying material reality that must be uncovered. From this view, exploitation is seen to exist insofar as material activity appears as something that it is not. By corollary, once we can reveal and bear witness to this misrepresentation, we are on the way to overcoming exploitation by it.

Marxian theories of algorithms involve another fundamental aspect, which is to conceptualize uses and operations of algorithms in terms of labor or economic value. While this conception enables Marx's entire theoretical apparatus to be brought to bear on algorithms, it comes at a loss of generality and specificity. For example, Christian Fuch's work on *digital labor theory* takes aim at social media platforms for obfuscating the value produced by data collected about user interactions, effectively alienating users from the fruits of their labor (Fuchs, 2014). Work in this vein sees online communities of users as marketable commodities and search engine results as rented property (ibid.), which users are not able to perceive as such.

On the one hand, these theories help to unpack the economic dimensions of algorithmic associations and orderings. On the other, they define various computational processes, effects, and appearances as value-creating, which motivates the idea that exploitation can be overcome by redistributing value in some way (e.g., by demanding an equal distribution of data, its use, or its profits). This restricts the use of these theories to algorithms that produce value when they are used, while also obscuring the effects and materiality of algorithms that cannot be conceptualized in terms of value production or extraction.

Other Marxian theorists of algorithms and perception retain fundamental ideas from Marx, but are less literal in their application. For their part, these theories are typically inspired in some way by Marx's writing in the Grundrisse, a manuscript that contains reflections on the role of technology in automating and mediating human activity (Marx, 2005). In contrast to Marxist theorists from the Frankfurt School, which tended to see in technology the materialization of a reductive scientific rationality, scholars who invoke the Grundrisse are interested in the ways that technology can shape – as well as emancipate – human thought and practice.

This includes the Autonomist Marxists, whose accounts of "immaterial labor" (Lazzarato, 1996) and "semiocapitalism" (Berardi, 2009) brought Marx's work to bear on the kinds of labor enabled by computer systems. But it also includes theorists who view algorithms as representational media that augment human intelligence, by inscribing memories (Stiegler, 2014), encoding social knowledge (Terranova, 2014), or shaping the meaning of communication (Jodi Dean, 2005). Work throughout this tendency conceptualizes algorithms like prostheses that augment and mediate human activity, perception, and knowledge. It is often informed by Marx's theory of a "general intellect," where human knowledge becomes materialized in technologies and technical practices, which can for that reason be appropriated toward emancipatory ends (Marx, 2005). This work is principally interested in how the symbolic systems and constraints introduced by communication technologies preclude emancipatory forms of perception, as well as how they might be redesigned or reconfigured to serve an emancipatory function nonetheless.

Work in this area has an affinity with Foucauldian and Deleuzo-Guattarian theories, which emphasize a concern with subjectivation: the process by which people come to perceive and moderate themselves, namely according to social practices and values that they perceive in their environment. In contrast to *subjection*, where classifications, norms, or regulations are imposed on individuals, Foucault's notion of *subjectivation* attends to the ways that people autonomously configure their behavior in response to their environment (). What is typically at stake for theories of subjectivation is that this autonomous activity is no freer from power; instead it become a site of power, where individuals engage in "techniques of the self" to condition themselves according to certain perceived values (). Foucault's theory of subjectivation, and his related notion of governmentality, theorize how power operates beyond outright subjection.

Most theories of the role of subjectivation in *computation* invoke Deleuze's notion of the "dividual" (Deleuze, 2017), which gestures at the idea that subjects can be moderated according to diverse and impersonal information about their behavior, as opposed to their discrete identities. With regard to algorithms specifically, Lazzarato follows Guattari to argue that computer technologies inaugurate a special type of "terminal subjectivity," characterized by people optimizing themselves as nodes in networks of information exchange (Lazzarato, 2014). A key feature of this type of subjectivity is its dependency on "a-signifying semiotics," or systems of meaning that are not codified like language, such as gestures and, perhaps, abstract data visualizations (Guattari, 1996). If algorithmic opacity addresses concerns about the invisibility of computational processes, Lazzarato's considerations about terminal subjectivity and a-signifying semiotics raise parallel concerns about how the visibility of these processes influences human behavior and self-perception.

Antoinette Rouvroy's theory of algorithmic governmentality (Rouvroy & Berns, 2013) approaches this concern through Foucault's theoretical lexicon, where governmentality names a mode of governance that operates through subjectivation, rather than outright coercion and subjection. But Rouvroy argues that subjectivation is mitigated by algorithmic systems that govern behavior without reference to visible social categories. In contrast to Lazzarato, who echoes Guattari's appreciation for symbolic systems that can escape the shackles of human determination and perception, Rouvroy's use of Foucault is more pessimistic about computation: the primary motor for human self-perception and reflection is undermined by the opacity of data collection and processing.

Other uses of Foucault focus less on establishing general theories of subjectivation than on conceptualizing agency in algorithmic governance; that is, how humans and algorithms interact upon one another, and what this entails for how we should regulate algorithms. For example, Lucas Introna's (2016) analysis of algorithms in terms of governmentality stresses the performative nature of algorithmic operations: they do not simply determine human activity, but are also enacted by humans. This complicates the idea that algorithms are agents that operate entirely outside of human behavior; instead their operations are implicated in human practices, and participate in the processes by which people perceive themselves. Similarly, Tobias Matzner calls for a conceptualization of subjects as distributed across algorithmic operations that they interact with, rather than autonomous actors who govern or are governed by algorithms (Matzner, 2016).

One contemporary theorist that explicitly combines Marxian and Foucauldian approaches in his theorization of technology design is Andrew Feenberg. Feenberg (2017) invokes Foucault for his insistence that power is implicated in every knowledge system, and therefore that there is no system of knowledge that guarantees a universal regime of equality free from the asymmetrical deployment of power. He also follows STS to analyze the values that inform the design of particular technologies, rather than theorizing technology as a totality that always reflects the same values and tendencies. Like STS, Feenberg argues that the values of technology design are imbued into their construction, which enables him to argue that resistance to technology design is also resistance to the cultural, political, and economic values that they operationalize. In the last instance, Feenberg's approach is also based on a Marxian dialectic: it rests on the idea that technology raises problems when it fails to meet the needs of individuals – that it is subjected to change and reinvention when it falls out of favor with them. This establishes a tension or contradiction between technology design and use that presupposes the possibility and utility of resistance, based on a failure of technology to meet individual needs.

Altogether, where Marxian work on computation differs most sharply from the other tendencies is in its estimation that algorithms are powerful tools limited fundamentally by the social uses imposed on them. While some scholars invoke Foucault to imagine how algorithmic systems could be designed to support subjectivation (see for example Bergen & Verbeek, 2021), the Marxian view traces the harms of algorithmic systems mainly to their configuration and ownership by capitalist regimes. From the Marxian vantage, algorithms are problematized on the basis of their *misrepresentations*, or discrepancies with reality, which can be revealed and corrected by exposing their material operations. Hybridizations of Deleuzian and Foucauldian tendencies are similarly concerned with the power afforded by algorithms, but moreso because they involve semiotics that influence subjectivation, rather than black boxes that conceal the truth of technological productivity and exploitation. Thus while Marxian theories favor an account of algorithms as technologies that materialize human intellect, and therefore might be redesigned and repurposed to humanitarian ends, others may see in algorithms a more pervasive capacity to influence human behavior, and particularly subjectivation, which is unlikely to be overcome definitively, outright, or according to a single program. Among the latter, work may be optimistic or pessimistic about whether this capacity of algorithms to influence subjectivation could benefit human lives.

Altogether, speculation about the limits and merits of computation tends to follow in line existing philosophical frameworks and concepts, such as subjectivation, the dividual, or a Marxian conception of labor. Others are based on traditions in media studies, like interface, rhetoric, and cinema studies, whereas others yet are broader in their scope, like algorithmic experience and semiotics. While recourse to these concepts can help us to understand the properties and behaviors of algorithms in terms of longstanding theoretical discourses, it also tends to constrain inquiry to the terms and principles of existing theories. Consequently, the actual capabilities and potential consequences of algorithmic technology appear to be spelled out in advance by existing theoretical ideas. This is one motivation for developing a phenomenological approach to algorithm analysis that, while drawing from existing empirical research and theoretical literature, is informed by specific properties and behaviors of algorithms, as well as by the practices involved in their design. In the next section, I develop a theory motivated by this idea.

Chapter 3: Toward a Theory of Paramediation

Theories of algorithmic governance face a challenge. On the one hand, conceptualizing algorithmic governance in terms of regulation, representation, and ontology precludes a nuanced understanding of interactive and phenomenological dimensions of algorithms, such as those involved in gamification. From this view, the harms of algorithms are attributed to their reifications of partial ontologies – that, is how they abstract the world according to profit-motivated heuristics (Jocque, 2022) or how they induce a subjective myopia that falls short of objective judgment (Smith, 2019). Accordingly, the resolution of algorithmic harms amounts to the improvement of ontologies: maximize inclusion in algorithm logic, make algorithmic abstractions more holistic, or establish regulations that compensate for their errors. Here, the goal to master algorithmic abstractions, to repurpose their mechanisms of representation to more humanitarian ends, begins to take precedence to a material analysis of what exactly algorithms do to subjects, and can do to them, beyond forcing society to adhere to the limitations of code.

On the other hand, conceptualizations of algorithmic governance that attend to governmentality and subjectivation fail to concretely specify the mechanisms or social practices that might be confronted to change the social consequences of particular algorithms. Theories of algorithmic imaginaries, experiences, superstructures, and governmentality, while essential for formulating a departure from a regulative model of algorithmic governance, remain unspecific about how algorithms are designed to interact with and govern human behavior. Indeed, these concepts may simply be used to extend an epistemology of algorithmic ontologies into new domains, where algorithms can be scrutinized for imposing ontological schemata on perceptions (Just & Latzer, 2017), affects (Bucher, 2017), visible information (Introna & Nissenbaum), or discourses (Beer, 2017). Even in more recent iterations of algorithm critique that aim to decenter a liberalist conception of autonomous rational subjects, we continue to see a notion of subjects regulated by algorithm logic and ontologies.

In this section, I propose an alternative to these perspectives: an attention to how algorithms are designed to arrange perceptible phenomena. While concepts like subjectivation enable us to acknowledge the capacity of algorithmic governance to operate beyond regulation, we still lack an understanding of how to confront this capacity in concrete approaches to criticism and design, beyond demanding better computational ontologies – or policies that respond to the consequences of the ones we have. The algorithmic arrangement of perceptible phenomena is a different register of algorithmic operations. Like algorithm logic and ontologies, it can be concretely analyzed as an object of design. But like theories of subjectivation and governmentality, it also accounts for the capacity of algorithms to influence subjects without simply imposing regulations on them. It can inform both of these perspectives by shifting attention to a particular register of algorithm design and operations, on the way to developing more concrete methods for evaluating their social consequences.

3.1 Paramediation

Algorithms sample information about material circumstances, arrange appearances according to these samples, and recursively coordinate or optimize these appearances to improve their effectiveness according to certain heuristics. Accordingly, I define paramediation as the algorithmic coordination of phenomena in dynamic relation to data. The connotation of "para-" is of something that happens alongside or parallel to something else – an appearance that is articulated according to other conditions encoded in data. Just so, algorithms can synthesize and arrange phenomena such that their appearance is fine-tuned according to a number of factors, or parameters, that operate in dynamic relation to one another.[2] What is made perceptible through paramediation is firstly derived from material circumstances encoded in data, and then designed to deviate from these circumstances in controlled ways; that is, parametrically. Ultimately, in a feedback loop, the resulting appearance may be calibrated according to data collected about its effects.

A simple example of paramedia is a mobile device notification: for example, when a mobile device vibrates to indicate to someone that they have received a message. What this vibration makes perceptible is determined by an algorithm that executes it according to data collected in real time. Previously, I used the term algorithmic semiotics to account for the fact that algorithms operate via percepts like this, and that attending to how they are designed is indispensable for understanding the social consequences of algorithms (Polack, 2020). A theory of paramediation, for its part, focuses on the way that algorithmic processes articulate data, semiotics, and material processes together to arrange appearances. Paramediation involves algorithmic semiotics, but points to the technical practice of arranging them, rather than their capacity for signification and their enlistment in the service of particular design goals (as in Polack, 2020). In particular, it points to the algorithmic logics of optimization and procedural generation through which phenomena emerge.

---

[2] Here paramediation relates to the word "parameter," which has different meanings throughout various disciplines, but generally refers to a discrete property that characterizes the behavior of a broader system. For example, in "parametric design," parameters are input variables which specify constraints, according to which a structure or process will be generated as an output. Parameters can thereby be adjusted to shape the output that results. As well, in machine learning, a hyperparameter is a parameter that is learned over time according to data input, rather than specified in advance. This type of parameter may be less legible to a human operator – there are typically thousands – but they are mediated and shaped by guiding heuristics. Paramediation implies both these types of parameters: those that are specified in advance by manual design, and also those that are generated through an algorithmic process. Both types of parameters are tuned and refined according to the outputs that they yield.

Here the use of the term "mediation" in paramediation refers to the process of conveying something to human reception through a designed technology. It positions algorithms as an intermediary between human perception and other practices, objects, and environments, which means that they also lend themselves to being used as *leverage* by those who design them, or points of dependency through which control can be exerted over dependents (Peters, 2015, pp18-21). Moreover, as theoretical traditions in communications and media studies have argued (Lievrouw, 2009), "mediation" is not limited to the autonomous agency of technology; it implicates the human interactions, interpretations, and social contexts that inform its reception and use. Mediation thus conceived is less of a one-directional transmission from technology to people as a process of articulation which is comprised of human and algorithmic agencies alike.

The status of data here is worth defining explicitly. Data, whose etymology points toward its status of being *given*, is exactly that for paramediation: something provided as an input which an algorithmic process uses to inform the arrangement of phenomena. The fact that data is given to paramediation does not necessarily mean that it is objectively true (Johanna Drucker's notion of data as *capta* points to the way that it is captured and arrested from material circumstances by design, rather than already existing in the world; Drucker, 2011). But data may nonetheless be accepted as *given* – in the sense of being an a priori truth – following a subjective or intersubjective appeal to its validity, or a specific procedure for establishing its truth. Paramediation leverages this quality of data's *givenness* to arrange phenomena. That is, irrespective of the reason that data is accepted as given, it arranges phenomena in accordance to what is given, so that it conforms or correlates to it. This means that paramediation fits in or correlates with the regime of truth that marks the status of particular data as given, and it leverages this truth to have an effect.

For example, beyond mobile device notifications, other examples of paramedia explored in this section include deepfake images designed to correspond to given images and maps of crime predictions designed to correspond to given data about crimes. In each case, paramediation ascertains a structure from data – which is itself regarded as given according to some truth procedure – and it arranges this structure to introduce corresponding possibilities for meaning-making and action. What makes paramediation so effective is that it leverages data about the given to arrange phenomena, making these resulting phenomena appear in harmony or correlation with the given. More significantly, it can leverage data about the effects that these phenomena have, for example, whether they provoke particular responses in subjects. Ultimately while mobile device notifications provide us with a trivial prototype of paramedia, the full implications of paramediation come into view when it involves the use of interaction and machine learning.

The first purpose of a theory of paramediation is to account for a technological capacity to govern behavior not through direct material coercion, but through the arrangement of appearances. Algorithms do not simply regulate activity or misrepresent reality; they arrange phenomena to evoke particular perceptions, which can be designed to influence activity in turn. And unlike virtual reality, paramedia are not created to replace other appearances, activities, and circumstances, but to operate *alongside* them. While other media – from the textual to the cinematic – are also designed to arrange appearances that emulate other perceptions and articulate fictional worlds, the breakthrough of algorithmic media is to arrange these perceptions according to collected data, which accord with phenomena that are already accepted as given, with respect to a particular regime of truth. This data-driven arrangement of phenomena enables something unprecedented by other media techniques: *parametric design*, or the establishment of constraints between interactions and statistical inference to synthesize emergent phenomena, which can then be optimized according to specific heuristics and data streams.

A preliminary question about paramediation is how generalizable the theory is. Is it applicable to all algorithms, or only some? Provisionally, we could specify some minimum criteria for paramediation. First, to reiterate, we are talking specifically about computer algorithms, which are encoded into hardware and executable: things happen in the material world when computer algorithms operate. Second, paramediation is limited to circumstances where algorithms are designed to involve or operate on human subjects. In other cases, algorithms can be designed to operate independently of any human awareness and involvement. In such cases, an attention to appearances and human perception in algorithm design is unlikely.

Third, paramediation is not necessarily involved when algorithms enact control over human activity absolutely. For instance, algorithms can physically lock a door to detain someone or automatically determine their insurance premiums – evading considerations about their perceptions entirely. Indeed, these deterministic algorithmic mechanisms are the main consternation of critical algorithm scholarship (even when the latter takes systems like social media platforms as its exemplar of algorithmic governance). Nonetheless, even in these cases, I argue that as soon as the effects of an algorithm can be sensed or felt, paramediation once again becomes an aspect of design. Therefore, in the case of deterministic algorithmic mechanisms, we can say that paramediation is not necessary but nonetheless implicated. Moreover, one concern of this thesis is that as deterministic algorithmic mechanisms come under fire for their determinations and biases, paramediation can serve as an alternative that apparently dispenses with these overt determinations.

The following subsections discuss how a theory of paramediation responds to issues raised in the previous sections: how the theory can be situated between phenomenology and epistemology, how paramediation relates to algorithm logic, and how the theory can be distinguished from principles of computer interaction design. Following this discussion, I conduct illustrative case studies to demonstrate how the theory applies to understanding some concrete instances of algorithm design and use.

3.1.1 Beyond computational ontology

A theory of paramediation directs attention to the concrete technical practices and mechanisms that enable it. To do so, it balances the limits of an epistemological approach (which disregards the effects of phenomenal arrangements that do not involve ontological assertions) with those of a phenomenological approach (which often stops at theorizing the social consequences of algorithms in terms of how people perceive them). To this end, paramediation names the material *process* designed to arrange phenomena, provoke sensations, and elicit particular interpretations and reactions, which can be concretely analyzed. In turn this process becomes apprehensible through paramedia, the phenomenal *result* of paramediation. This double nature of paramediation – the way that it must be apprehended as involving both a material process and a phenomenal outcome – is critical for grasping its mode of production, and how its consequences escape the purview of either epistemology or phenomenology alone.

Neither perceptions of algorithmic outcomes, nor the materiality of algorithm processes, provides us with a complete account of paramediation. While people may differ in their interpretations of these phenomena (phenomenology of paramedia), the material processes that produce them exist irrespective of observation (materiality of paramediation). Thus we must attend to what algorithms are designed to make perceptible – what they are designed to mediate and evoke – which is a materialist project as much as a phenomenological one.

Here we have recourse to a *post-phenomenological* conceptualization of technology, coined by the philosopher of technology Don Ihde (Ihde, 2008). Post-phenomenology attempts to resolve an incompatibility between traditions in phenomenology and materialism, which Ihde attributes to phenomenology's overemphasis on subjectivism. Ihde insists on distinguishing essential tenets of phenomenology, like its acknowledgement that *interpretations vary* and its appeal to *embodiment* against a transcendental faculty of mind, from appeals to subjectivism, which he reads as a dated strategy for contesting the subject/object dualism of Cartesian epistemology. By distancing itself from subjectivism while retaining other considerations of phenomenology, post-phenomenology aims to broker between technological determinism, which holds that technology shapes the interpretations and actions of subjects, and social determinism, where interpretations and uses of technology inform its consequences. For post-phenomenology, the materiality of a given technology and its interpretations entangle to produce a *multistability*, consequences that are simultaneously varied and constrained.

A theory of paramediation takes after post-phenomenology to find a balance between materialist and subjectivist interpretations of algorithms. What it retains from theories of algorithmic materiality, and more broadly from appeals to examining algorithm code, logic, and software, is an acknowledgement that algorithms are designed in particular ways to have particular effects on subjects, and that they can inform how people tend to perceive the world. Nonetheless, this influence is seldom deterministic, even though it may be hard-coded into algorithm logic. What the theory of paramediation brings from phenomenology — represented in algorithm studies today by research on "algorithmic imaginaries" (Bucher, 2017) and "algorithmic experience" (Rossiter & Zehle, 2015; Alvarado & Waern, 2018) — is its insistence that the consequences of algorithms are not only irreducible to this logic, but also increasingly unlikely to be found there. We miss the consequences of paramedia when we are exclusively concerned with the bias, transparency, and representations of algorithm logic.

A theory of paramediation shifts from an epistemological analysis of algorithm ontologies to a post-phenomenological analysis by addressing how algorithms are designed to manifest to human perception, which in turn motivates a shift in our understanding of what it means to be governed by algorithms. Instead of viewing algorithms as procedures that aim to accurately represent the world, govern activity, or extend human abilities through networked devices and sensors, we attend to how algorithms govern behavior through the arrangement of appearances. This is not to argue that algorithms do not make ontological assertions; just that their consequences are irreducible to them. This, not only because algorithmic systems exist which do not involve ontological declarations whatsoever, but also because those that do operationalize them in ways that cannot be exhaustively understood in terms of their ontological claims; they act and appear in ways that do not concern assertions about what exists or ought to exist.

This means that our understanding of algorithmic governance should not be reduced to the capacity of algorithms to regulate action by enforcing opaque rules, norms, or (mis)representations of reality. Algorithmic governance also involves paramediation: the exhibition of possibilities for action and interpretation informed by procedural interactions among rules and human activity, whatever ontological assertions or inferred norms they may involve.

In contrast to an epistemological approach to algorithm analysis that critiques algorithmic ontologies, a post-phenomenological approach is concerned primarily with the relationship between computation and its consequences — not with that between computation and its representation of reality. This does not undermine a concern with the provenance and epistemic validity of data, but to the contrary, broadens our understanding of algorithm consequences to include phenomenal consequences that escape an epistemological purview. An implication of this approach is that algorithmic ontologies are themselves perceptible as phenomena, designed to be perceived or enacted, and for this reason irreducible to their own terms. The computational representation of reality cannot be understood as solely ontological; it has phenomenal consequences that have naught to do with a congruity between computational ontologies and the objective reality they suppose to grasp. For this reason, the effects of algorithm logic cannot be adequately accounted for by algorithm logic alone, without attending to their sensible manifestations and the technical practices developed to evoke them.

3.1.2 Beyond logic

Altogether, modern computation operates at a level of abstraction that tends to be intractable and impractical to attribute to particular components of algorithm logic and ontologies. Neural networks, examined in one of the case studies in this section, provide a clear example of this because their complexity precludes reverse engineering precisely how their parameters result in particular actions. Their parameters, or hyperparameters, are learned over time through exposure to various examples, rather than programmed in advance. The media theorist and algorithm scholar Luciana Parisi has argues that the semi-random processes used to arrive at these parameters point to a fundamental "incomputability" in algorithm operations (Parisi, 2016). For Parisi, central to the epistemology of machine learning is its dependency on indeterminacy, as opposed to a definitive specification of logic or ontology. Machine learning cannot be adequately thought as the mere automation of data inputs and hardcoded ontologies — incomputability defies a conceptualization of machine learning as the direct translation of ontology into action.

But even beyond machine learning, simple digital images have effects that are irreducible to the code that instantiates them: how a digital image is interpreted, and what it signifies to someone, is something other than the data, display technology, and compression algorithms that produce it. Indeed, while we take for granted that human interpretations of photographs or illustrations can vary, and that we depend on our subjective perception to understand them, the fact that digital percepts are generated by computational rules deceives us into thinking that the effects of algorithms can be traced back to their logic in all circumstances. We err by trying to think such phenomenal effects of algorithms exclusively in terms of their pre-programmed logic. Not only the semi-random incomputability of algorithm logic – but also the phenomena that algorithmic logic evokes in even the most banal circumstances – cannot be understood by way of code or ontology alone. While advancements in deep learning are becoming able to associate images with textual descriptions, as if they have decoded percepts as ontologies, what these images and their textual pairs signify to people is irreducible to the algorithm logic and data that inform them.[3]

[3] Some scholars also advocate the use of symbol manipulation in machine learning, which they argue would make deep learning legible in terms of discrete symbols rather than hyperparameters (for example Marcus, 2022). However, insofar as learning is not absolutely discrete and involves the configuration of perceptible phenomena (as in image-to-image translation, for instance), this would only abstract hyperparameters in a new way that introduces new kinds of obfuscations.

This suggests that we cannot grasp the consequences of algorithms without attending to the phenomena that they are designed to configure, generate, and operationalize. Undoubtedly, algorithm logic produces phenomena, and programming languages are used to abstract computational processes that produce phenomena, so it would be wrong to say that attending to phenomena is something totally distinct from attending to logic and code. But again, we err to think that algorithmic phenomena can be adequately apprehended through algorithmic logic alone. A trivial example is the `printf()` function in Java, which is used to display characters on the screen when the code is executed. While the universal accessibility and popularity of this function may make it appear to symbolize the procedural display of characters unequivocally, imagining such a simple correspondence between code and appearance neglects the architecture of screen display hardware, software, and the programming library designed to make this possible.

But the trivial example of `printf()` does not address what is really at stake in paramediation: the emergent phenomena produced by more complex algorithms, which I examine in the next section. For these algorithms, the tendency to explain their consequences in terms of algorithm logic reveals a great gap between computational ontologies and the phenomena they evoke. Appeals to an epistemological view of algorithms ignore this gap, when they claim that an attention to algorithm ontologies can supersede ways of knowing algorithms that depend on understanding how they appear – normally raised by subjective experiences and perceptions of algorithms. Against subjective concerns that these algorithms can cause harm, system developers can insist on the rigor of algorithm logic and ontologies (Polack, 2019), even if it exists at a gap remove from their phenomenal consequences. This reinforces a regime of exclusion from decision-making about algorithms, founded on distinctions between objective and subjective knowledge about algorithms, treated as technical literacy and illiteracy respectively (ibid.).

Here the intervention of a theory of paramediation, then, is to identify a domain of perceptibility that enables us to understand the social consequences of algorithms, other than ontological representations instantiated in code. Far from a disavowal of computational materiality, this is a recognition of how this materiality is designed to systematically arrange phenomena and elicit perceptions. A key advantage of this theory is that, while it centers the importance of lived experiences of algorithms for understanding their consequences, it also acknowledges the materiality of these experiences: aspects of algorithmic operations and appearances that – while distributed across algorithm logic, computer devices, and various phenomena in space and time – can be concretely apprehended and analyzed. Without forgoing an insistence on the materiality of algorithms, this conceptual shift should develop our understanding of how algorithms and subjects are agential in interaction with one another.

3.1.3 Beyond interaction

This conceptual shift also motivates a change in our understanding of interaction, away from how it is normally conceived in computer sciences. Traditionally, user interface design (UI), human-computer interaction (HCI), and theories of algorithmic transparency have maintained that interfaces are supposed to provide useful representations of computational operations. This implies a process of communication between people and computational agents that interface design should optimize to ensure intelligibility, either by simplifying its complexity or emphasizing important aspects to enhance navigability. Graphical user interfaces are seen to reveal a subset of computational information to users, such that users obtain an adequate representation of how they can interact with this information, and what the information reveals.

A theory of paramediation acknowledges interaction as something more than a bidirectional exchange between people and computers where possibilities for action are prescribed and knowable in advance. Interactivity is not simply choice among computational rules and their pre-programmed responses, but the coordination of distributed processes that produce spatiotemporal phenomena capable of evoking affects like surprise, progress, repetition, anger, fear, anxiety, attention, or coherence. Users may participate in the interactions that produce these phenomena, or not; they may simply be on the receiving end of interactions among algorithms and data, or what we might call *para-actions*. A theory of paramediation is concerned with how user interactions and algorithmic para-actions are designed together such that phenomena emerge to users in specific ways.

The design of user interactions and algorithmic para-actions is an important aspect of paramediation that enables events and significations to emerge dynamically without needing to be pre-programmed explicitly into algorithm logic. If an epistemological view of interactivity defines interactions with algorithms as what their logic makes visible, accessible, and thereby *usable*, a phenomenological view would find a more adequate model of interactivity in *play*: interactivity is the ability for actions to emerge among constraints without their being prescribed in advance, where determined rules enable actions to *play out* in a way that is only partially predictable. From the vantage of paramediation, algorithm design is not just the specification of procedural rules that regulate or model reality, but the strategic architecting of perceptible phenomena. Rather than regulating activity, paramedia produce possibilities for meaning-making and action.

Rather than contradicting or misrepresenting facts given in data, algorithms calculate relationships between them, encoding them into ontologies or inferring their norms, and they arrange phenomena to be situated among or alongside these norms. These interventions of algorithms are perceptible to people as something that corresponds or deviates from normal circumstances in a coherent or patterned way; something *sensed* before it can be definitively rationalized or traced to a particular agency. This is because, rather than appearing to *replace* given phenomena (virtual reality) or *supplement* them (augmented reality), paramediation *arranges* new phenomena so that they mix in with or *correlate* to given phenomena, according to some heuristic that measures the pragmatic effectiveness of this correlation.

3.2 Illustrative Case Studies

In this section, I use case studies of four algorithmic systems to illustrate and elaborate the theory of paramediation. Each case involves a specific application of algorithms that raises unique considerations: image-to-image translation, content recommendation and incentive design, data analysis for crime investigation, and procedural generation. Each case can also be seen to involve a certain type of media: image, event, narrative, and world. For each case, I describe how the algorithmic system works, some of the ethical considerations it raises, and what a theory of paramediation can enable us to understand about these considerations. I conclude the section by identifying some aspects of paramediation reflected throughout the cases.

3.2.1 Image-to-Image Translation

Image-to-image translation (I2I) is the practice of translating visual features of one image to another. Some of its popular applications include converting a blurry image to a high-resolution image (super resolution), converting an image to a 3D image (3D object generation), producing images of a person's full face provided an image of their face's profile (face frontalization), using line drawings to synthesize photorealistic images (conditional image synthesis), reproducing images with the aesthetics of famous painting styles like impressionism (style transfer), and replacing faces in images or videos (face swapping).

I2I is an intuitive example of paramediation: more conspicuously than other algorithmic techniques, it shows how algorithms can arrange phenomena according to various data to synthesize new percepts. I2I also raises clear concerns about the ability of algorithms to misrepresent the truth, often described as the production of "deepfakes," or algorithmically generated depictions of counterfactual circumstances. The capacity of I2I to manipulate the truth has motivated a widespread interest in ensuring that the effects of algorithms correspond to the truth accurately, rather than misrepresenting it.

From the vantage of paramediation, however, this interest in representation obscures a different capability of algorithms: the ability to produce phenomena that evoke particular perceptions. This capacity cannot be adequately apprehended by an epistemological analysis of algorithm ontologies, which is concerned with whether algorithmic phenomena are objective. Instead it requires an attention to what algorithms are able to make perceptible irrespective of their objectivity. Here I describe fundamental principles of I2I, and then examine its implications for a theory of paramediation.

3.2.1.1 How I2I works

I2I involves designing algorithms with the explicit purpose of producing perceptible phenomena. This process typically involves compiling a set of reference images, or "source data," and a corresponding set of data that provides an example of how the true data images should be transformed, called "target data." I2I algorithms are designed leverage the target data to produce distortions of the source data, or "candidates," which retain some aesthetic or semiotic characteristics of both, but introduce specific variations. To accomplish this, visual patterns or "features" of the source data must be identified, and subsequently encoded into a computational format, such that they can be used to reconstruct images with some variations.

One state of the art approach to I2I is to train an artificial neural network (ANN, or neural net) to produce randomly generated distortions of the source data, or "candidates," that can be compared to the ground truth of the source data. To enable comparison, ground truth target images and algorithmically generated candidate images are randomly mixed together and used as input into a second neural network, which has the goal to identify which of the images came from the true data distribution (target data), and which were generated (candidates). The combination of these two neural networks, the candidate "generator" and the true data "discriminator", is called a Generative Adversarial Network (GAN).

In the last year, a different approach to I2I has soared in popularity, which makes use of transformers instead of GANs. Transformers are models that break input data instances (e.g., images, sentences) into individual elements (e.g., pixels, words), which can be inspected one by one to identify their interrelations and relative significance in each input data instance. By ingesting massive amounts of data and storing relational information about these individual elements, transformers have proven to be extremely effective at predicting patterns, provided some individual element as a prompt. One famous example is GPT-3, a natural language processing model trained on billions of example sentences that can respond to human prompts organically. Google Translate is another. In I2I, transformers can be used to identify simplified patterns in images, as well as their correlations with patterns in text, like captions that describe what is depicted in images. These identified patterns can then be used as a reference point to encode images into a simplified computational representation, which can subsequently be used to reconstruct the same image. Because patterns in images and text, and their correlations, inform this encoding, a new image or text can be used as prompt to respond with a corresponding image.

For either GANs or transformers, training algorithms in I2I means using large data sets to extrapolate patterns from images. By breaking images down into parts, and identifying whether random permutations among these parts resemble the original images (GAN method), or otherwise calculating how these parts relate to others (transformer method), statistical patterns in visual data can be extrapolated that are not readily visible to the naked eye. These patterns reflect statistical norms according to which new images can be synthesized. Whereas the true data distribution follows a norm encoded in data, the distribution of the image that results establishes a complementary norm in nonlinear relation to the original, which I call a *paranorm*. This enables synthetic images to cohere with normal phenomena while introducing counterfactual, paranormal novelties that respond to certain I2I tasks.

As a whole, the improvement of these techniques evidences a developing technical capacity to manage the arrangement of perceptible phenomena. This involves a developing specification of problems, criteria, and solutions for manipulating percepts in specific ways. It includes the identification of algorithmic methods that can display certain phenomena, statistical techniques that are generalizable across problems, data processing methods that distort input images in ways that are reproducible, coherent, and generative, heuristics for evaluating generated images, best practices for responding to undesirable arrangements of phenomena that emerge as artifacts from algorithmic processing, and types of input data that improve or compromise the consistency of these arrangements. These practices all reflect developments in the design of paramediation: algorithms become more readily able to identify various types of statistical norms in data, and subsequently able to distort and arrange them in highly tailored ways. The phenomena evoked as a result of this process appear to emerge from real circumstances, but exhibit a quality that deviates from them significantly, subtly, or nearly imperceptibly.

The function of paramediation in I2I is thus to balance between the norm of a true data distribution and the abnormal determinations of algorithmic manipulation to arrive at a paranorm that manifests novel synthesized images. At either side of this balance are extremes that lead to failure. On the one hand, synthesized images can appear too similar to source data, essentially reproducing them without variation – they are too normal. On the other, they may appear sufficiently different, but so distinct from the source data that they are completely incoherent and fail to realize an I2I task – they are too abnormal. One way that this balance can manifest is as the *algorithmic paranormal*, or an appearance that appears uncanny to people in a way that betrays its artificiality – it may be either too normal, too abnormal, or something in between. I2I must carefully balance between the normal and the abnormal in establishing a paranorm, so as to not fall into either extreme.

Critically, the difference between normal and paranormal is not always apprehensible in terms of the statistical distribution of data, which I2I algorithms may use to evaluate the similarity of synthesized images to target data. In such cases, the relationship between normal and paranormal might be measured in other ways, namely by soliciting feedback from human perception. For example, the Learned Perceptual Image Patch Similarity (LPIPS) metric uses crowdsourced data from human workers on Amazon Mechanical Turk to train neural networks to calculate image similarity (Zhang et al., 2018). This is one measurement of the relationship between normal ground truth images and paranormal evoked images, among others.

3.2.1.2 I2I and Paramediation

Theoretical accounts of I2I tend to emphasize that the images they produce defy the truth in a way that is unprecedented by traditional, non-algorithmic media. These theories tend to focus on "deepfakes," or algorithmically synthesized media that depict counterfactual circumstances (e.g., a real person depicted in an untrue circumstance). Deepfakes are said to contribute to a "post-truth" regime (see for example Antinori, 2019; Chesney & Citron, 2019; Hasen, 2019), since they produce depictions of entities and events that do not exist, but which nonetheless appear to. The capacity of algorithms to deceive human perception with synthesized images in this way raises a number of concerns in this scholarship that might be briefly summarized as follows:

1. evidentiary problem: algorithms can produce convincing media that contradict the truth, compromising the reliability of media as documentary evidence
2. ethical problem: people can use algorithms to produce media about others that portray them in deceptive, disparaging, and altogether untruthful ways, and it is unclear on what grounds the algorithm user or designer could be held responsible for this
3. epistemic problem: algorithms can produce convincing and believable media, without requiring an explicit specification of the truth to be programmed into their logic

Each of these problems rests on an ontological conceptualization of the truth, whereby an image is understood to be true if it depicts entities and events that exist. A variety of solutions have been proposed to correct this discrepancy between digital media and the truth, including computational methods for detecting deepfakes and for tracking the provenance of source material to verify its fidelity.

Despite the importance of this work in specific domains, it is unlikely to eliminate the existence of deepfakes or obsolesce their social impacts. To do so, deepfakes would need to be detectable, either by trained observers or automatically, such that they could be censored or labelled as such. This remains unlikely because the capacity of algorithms to detect fake images tends to develop at the same rate as the capacity of other algorithms to generate them. In particular, deepfakes can be disseminated to evade these algorithms, distorted with techniques designed precisely to fool them, or even used to produce true images to compromise detection accuracy.

This means that, so long as there are algorithms that can arrange perceptible phenomena, effective algorithmic manipulations of the truth are here to stay. But for a theory of paramediation, these manipulations would not be considered the exception of algorithmic operations – an improper use case to be detected and eliminated – but the norm. From the vantage of paramediation, algorithms are always designed to produce phenomena that cohere with the given and evoke something new about it; it is simply in the case of deepfakes that this newness appears flagrantly counterfactual when exposed. In such cases, the process of paramediation fails, revealing the *algorithmic paranormal*, or uncanny synthetic figurations that betray their artificiality.

What we are witnessing with I2I is a growing capacity to arrange phenomena in counterfactual, yet remarkably coherent ways. The task to identify which digital images are fakes – to ascertain whether they deviate from existing facts – is only readily applicable in the most extreme cases, where I2I is used with a specific intention to contradict the truth. In all other cases, we must identify more generally how algorithmic arrangements of phenomena can evoke particular appearances, beyond (mis)representing the truth.

For paramediation, instead of evaluating deepfakes, I2I technologies, and algorithms in terms of their claims to the truth, we examine more specifically how phenomena are coordinated to influence human perception, and to what effects. Fundamentally, this enables us to acknowledge that I2I media has effects that can cause social consequences and harm irrespective of their truth value. For example, the use of face swapping algorithms to superimpose celebrity faces onto pornography is not simply a problem because it contradicts the truth, but namely because of how it causes someone to be perceived. The truth of these synthesized images does not need to be in question for their harms to be felt, and deepfake media that is clearly marked as such can still evoke such an affective response – in such cases its correspondence to the truth does not influence its reception. This recalls the "paradox of fiction" in literature: fiction can evoke a real emotional response even if it does not presume to represent the truth.

To better understand the social consequences of I2I algorithms, we might attend to the algorithmic configuration of paramedia, which are neither totally true nor totally false. This approach concerns different questions about the ethics, consequences, and epistemology of algorithms: rather than questioning whether the content of an image is true, we interrogate how the content of the image was made to be perceived, to what ends, by whom, for whom, and under what circumstances. To answer these questions, we must go beyond detecting and regulating the use of I2I algorithm logic, to examine how people perceive digital media and its algorithmic synthesis.

These considerations are being examined by programs like the MIT Media Lab's Detect Fakes project (Groh et al., 2022), which produces educational materials that instruct users how to discern deepfakes, and collects data about their estimations to determine best methods for deepfake detection. While Detect Fake's research specifically examines whether humans or algorithms are better at detecting real images, it also motivates research in developing techniques that prompt users to think about the accuracy of digital images they encounter online. Such approaches move beyond the epistemological purview of fact checking and computational ontology to include more contextual, phenomenological considerations, including how images are perceived differently by users depending on the information they are exposed to and the systems they interact with. As opposed to attempting to bring the truth of images in line with the truth of their subject matter, the Detect Fakes project holds the implications and artifacts of visual phenomena are to more robust, context-sensitive criteria of evaluation.

An analytical shift from epistemology to phenomenology may seem trivial in the case of I2I; after all, the explicit purpose of I2I is to produce phenomena, not to impose ontological schemata on the world. Nonetheless, the evaluation of I2I's social and practical consequences in terms of epistemology remains common, as critical research and artwork continues to focus on how I2I misrepresents and miscalculates the truth. Altogether, the claim of a theory of paramediation is that a shift from epistemology to phenomenology is important for understanding other uses of algorithms as well – especially those that are commonly analyzed in ontological terms.

3.2.2 TikTok

Social media news feeds are a popular object of research for algorithm studies. They provide a familiar example of algorithms and their social impacts, and raise a variety of concerns about the use of algorithms to exploit human attention for profit, to limit access to information in order to optimize attention, and to do so without human awareness. In this section I analyze TikTok because the social media platform is known for its effective algorithmic content recommendations, which are designed to immerse users in an infinite stream of video content. These carefully designed automations have raised a number of specific concerns about the TikTok algorithms, including their promotion of harmful content and addictions, their incentivization of competitive labor, their systematic censorship of social movement content, and their racial biases.

TikTok is also an interesting case for investigating paramediation because, like most social media platforms, its algorithms are not absolutely knowable – its effects cannot be definitively traced back to algorithm logic at all. This partially knowable system could be called 'gray boxed' because it partially conceals the operations of algorithms, rather than concealing them absolutely (i.e., as a black boxed algorithm does). But the gray box metaphor is deceptive because all algorithmic systems exhibit some kind of partial visibility, insofar as they have effects on human subjects. Moreover, the complexity and stochasticity of machine learning techniques ensures that algorithm operations such as those used in social media are never absolutely transparent.

Because of the partial concealment of TikTok algorithms, our best approach to understanding them is to reverse engineer them: interacting with algorithmic processes and observing their effects, sometimes in combination with public documentation or patents that describe their design features. Significantly, the most knowledge we have about TikTok algorithms comes not from researchers, but from users. Here paramedia on the TikTok platform become objects of analysis in their own right, as the conditions of perceptibility and arrangements of appearances that make the TikTok algorithms agential become apprehensible to users.

3.2.2.1 How TikTok works

The #ForYou feed is the first thing users see when they open the TikTok app on their mobile devices. Even for new users, a stream of posted video content begins playing without being prompted. New content is introduced to users by a recommendation algorithm, also called a recommendation engine or a recommender system. Initially, the recommendation algorithm exposes users to content according to categories of interest that they select when they create an account, as well as their language, location, device settings, and data retrieved from third-party social media accounts used to log in. At this stage, content is recommended according to an existing norm of user activity associated with these generic preferences and properties. As the user scrolls through the #ForYou feed, further content is recommended according to the accounts that users follow, the posts that they like, and what content they pay attention to. This information comprises the user's interests or "interest characteristics" – the norm that their content is organized around, which is flexible and dynamically updated.

Indicators of user interests include the amount of time a user spends watching a post, especially to completion, and their interactions with posts. When a user exhibits a measurably high amount of interest in a particular post, the recommendation algorithm includes similar content on their feed (content-based filtering), and also recommends that post to other users who pay attention to similar content (user-based filtering). The algorithmic association of content and users to calculate recommendations is called collaborative filtering. Both are based on norms derived from data, not unlike the I2I algorithms that derive statistical patterns from image data.

Collaborative filtering on TikTok uses many algorithmic techniques to group users according to the posts and accounts they watch, interact with, and follow. Zhengwei Zhao proposes an explanation of these algorithms according to patents and public documentation from ByteDance, which operates TikTok, and a report from Cao Huanhuan (Zhao, 2021):

- **Partitioned Data Buckets** When new content is posted to TikTok, it is recommended to a random, relatively small group of users – a "bucket". If all of these users exhibit an interest in this content, the content is then recommended to a random, relatively larger group of users. The content propagates 'up' to larger groups of users the more that these random users show interest, and it also propagates back 'down' to smaller groups of users to test their interest. Through this process, the types of users that are interested in the post and its characteristics can be identified.
- **A/B Testing** Within buckets, a fraction of users are randomly selected and treated as an experimental group, whereas the rest remain as a control group. The experimental group is exposed to particular kinds of recommended content, and the interest level of

users in this group is compared to users in the control group. Multiple tests can be performed on this same experimental group to identify relationships between user and content characteristics.

- **Interest Label Matching** Users and content are labeled with interest tags, provisionally according to interests and hashtags they provide, and ultimately according to content they watch. Each user has an interest set compromising multiple interests. If interests A and B frequently appear in many users' interest sets (i.e., both A and B), at the same time as they less frequently appear alone (i.e., either A or B), these interests are considered relevant to one another. Thus users with one of these interests will commonly be recommended content associated with the other.
- **Time-Sensitive Learning** Content is tailored according to the time and duration that users use the app; for example, shorter videos may be recommended to users who are online for shorter periods of time.
- **Location Clustering** All users' location data is clustered (i.e., locations that are near each other are grouped together according to their overall geographic distribution), and user location is defined as the center point of the nearest cluster, such that they are recommended content from around this region.
- **Online Learning** The recommender system updates its model of user preferences in real time.

TikTok has also explained that characteristics of content are identified by hashtags, captions, visual effects, user interactions, and post sounds (a staple of the TikTok platform is that users select music and sounds for their posts) associated with it.

3.2.2.2 TikTok and Paramediation

The purpose of recommending content by similarity is mainly to attract attention from users, thus ensuring that they spend more time on the app. This increases the time that they can be exposed to advertisements, and it increases the amount of people posting and interacting with content, facilitating a more attractive platform for new users and enabling the more widespread collection of user data. This is the first instance of paramediation of the TikTok platform: the recommendation of media content such that user activity remains high, which is the arrangement of particular phenomena to optimize a particular behavior.

Here paramediation rests on calculating various norms about user behavior and content relationships, and then using these norms to tailor content to each user. Like paramediation in I2I, the goal is to retain elements of an existing normal distribution, but also to modify them to novel ends, according to a specific task. For its part, TikTok introduces content to users that falls somewhat outside of their typical recommendations, so as to introduce more varied content, and to collect more data about what kinds of content receive attention from users. By exposing users to more diverse content, models of their attention and preferences can be further refined. In this way, paramediation develops flexibly, introducing novelty in specific ways without deviating too far from the norm of user interest characteristics.

While the criteria and motivations for recommending content have been documented publicly by TikTok, further nuances of the recommender system are speculated by users and researchers who apprehend the algorithm's recommendations and attempt to reverse engineer them. Indeed, a key aspect of paramediation is that, since it involves the reproduction of normal patterns as well as the synthesis of patterns that deviate from it, it is readily perceptible to users as patterns of similarity and difference, consistency, and inconsistency. For example, users can apprehend the systematic paramediation of content recommendations by noticing the regularity or idiosyncrasies of content that is displayed to them on their #ForYou feed. A system designed to evoke particular perceptions from users – optimizing their engagement – becomes apprehensible to them through its operations.

Another kind of paramediation is at play in the semiotics of "exposure." Exposure indicates to users how their own content is displayed on other users' feeds. It includes post view counts and notifications of user interactions, which indicate to users when their content is receiving more attention. One of the purposes of exposure is to demonstrate to users which characteristics of their content are more attended to by other users, and ultimately to incentivize them to produce more content with these characteristics. If paramediation in the recommender system evokes the association and repetition of certain patterns in content, the paramediation of exposure evokes an association between the type of content that users post and its popularity. Here, the norm from which the paranorm is derived is not a statistical distribution of inferred patterns as in the case of I2I and the TikTok feed, but a causal relationship between a user's activity and their popularity, which they can infer themselves. Paramediation comes into play when algorithmic rules that determine the exposure of user content become apprehensible as a deviation from this norm. A sudden hike in user exposure and popularity is a paramediated event, which signifies to the user that they have done something right.

This causal paramediation involved in exposure is more able to influence user behavior the more it (1) varies (2) perceptibly (3) and consistently over time. If exposure is constant irrespective of user activity (too normal), users will not be incentivized to try new things and prioritize novel behaviors to optimize their exposure. Instead, to influence user behavior, phenomena must be arranged to evoke a tacit correlation between user behavior and exposure, such that certain activity yields a certain response. This response should vary over time, and namely increase, such that varied interactions with the app over time generally appear to cause increased exposure and a greater benefit to users. By enabling users to perceive how their activity corresponds to their exposure, paramediation motivates particular user behaviors. Exposure leverages paranormal deviations from a norm of user behavior to evoke certain reactions.

Although exposure is normally mediated through various notifications and metrics, an interface view in TikTok's "Creator tools" menu called "Analytics" represents exposure as a series of histograms and graphs. This interface gives users the option to select specific time periods to see how their views and followers changed over time in relation to their content. In this way, users can monitor their exposure more precisely and identify how its fluctuations correspond to their activity. While TikTok's documentation discourages users from interpreting fluctuations in their exposure as indicators of their performance,[4] committed users nonetheless depend on analytics to optimize engagement with their content. When a user interprets these analytics as indicators of best behaviors, the influence of paramediation on behavior grows.

Social media algorithms optimize user engagement by identifying correlations between their behavior and what is displayed to them. In many cases, these correlations result in content recommendations that are offensive, harmful, or otherwise disturbing to users. On TikTok, these issues are raised with regard to the promotion of content that depicts disordered eating, for example. TikTok's recommender algorithms systematically promote behaviors with disordered eating to users who attend to them most, reinforcing harmful behaviors. Users who have eating disorders and attempt to avoid this content can also be exposed to it by the recommender algorithm in the future, since the algorithm ultimately decides what content is displayed.

Other users and researchers have raised concerns that TikTok's algorithms exhibit racial biases (Strapagiel, 2020), as well as systematic preferences toward more attractive people.[5] The timing or accuracy of recommendations can also be harmful in itself; for example, a user could be upset by content that depicts something related to death when they have recently experienced death in their lives. Or more commonly, a user could feel that recommended content is so relevant to their lives that its accuracy is creepy, leading them to feel uncomfortable about their use of the app altogether.

Systematic approaches to addressing the harms of algorithms on social media typically fall into two categories: legal and technical. Legal approaches establish social media users as legal subjects with particular protections against abuses of their data. Technical approaches are

[4] https://www.tiktok.com/creators/creator-portal/en-us/tiktok-content-strategy/understanding-your-analytics/

[5] https://www.youtube.com/watch?v=1nAoIRBLzoc&t=720s

concerned with improving algorithmic systems so that they do not cause particular harms in the first place, such as by improving recommender systems to identify posts that are "appropriate" for users with greater accuracy. Both approaches tend to limit the regulation of algorithmic systems to identifying where specific automated decisions cause harm, which overlooks harms of algorithmic systems that cannot be conceptualized as violations of data privacy, use, or non-alignment between user interests and content properties.

Legal approaches are reflected by Terms of Service Agreements that specify what social media platforms can do with data, which excludes any mention of the various effects that algorithms can have on user perception and experience. This disregards the harms of, for example, exposure to content that encourages harmful behaviors for particular users but nonetheless satisfies Terms of Service agreements, algorithmic processes that imply associations among users with particular races or identities, and the automatic censorship of political content. Legal interventions operate primarily by incentivizing social media platforms to handle user data with greater security and scrutiny, and by using existing legal conventions to inform the moderation of user activity. It is rare for these interventions to cause changes to the logic of machine learning or recommendation algorithms, since the perceived effects of algorithms or data misuse are not exclusively a function of algorithm logic.

Among technical approaches, research is principally concerned with improving the "quality" of recommender systems, which is quantified by testing the systems with users in online trials (experiments where real users interact with the system to test it). Many methods and metrics exist to evaluate the quality of recommender systems, such as whether users "trust" the recommender system or feel that its recommendations are "novel" (Shani & Gunawardana, 2011; Avazpour et al., 2014). These methods may also attempt to develop heuristics for approximating these user perceptions with requiring their input, in offline trials (experiments where synthesized "dummy data" simulates the behavior of users to test the system). "Novelty," for example, can be approximated by measuring the similarity between recommended items, according to some criteria – another statistical formulation of the normal and the paranormal (ibid.). "Risk" has been approximated by measuring whether the algorithm tends to return incorrect recommendations (i.e., false positives; ibid.).

An attention to paramediation in TikTok acknowledges how recommender algorithms are designed to produce phenomena that elicit particular user experiences and namely behaviors. Of course, this capacity of social media algorithms to evoke user responses is widely acknowledged and critiqued, according to such ideas as "the attention economy" (for example Crogan & Kinsley, 2012) and "dark patterns" (Brignull et al., 2015) which refer to profit-motivated decisions to design algorithms and user interfaces that trap or deceive users into taking up certain behaviors. Interestingly, whereas harmful effects of paramediation are called dark patterns to highlight the transgressions of social media platforms and their designers, these same paramedia may be classified as bearing a 'poor algorithmic experience' from the perspective of working to improve these systems and mitigate their harms (Alvarado & Waern, 2018). Both of these ideas reflect an interest in how algorithmic systems are designed to inform human behavior and perception.

For legal and technical approaches, harmful effects of algorithms are typically qualified as aberrant, so that they can be highlighted, penalized, and ideally eliminated. This practice tends to portray algorithm design as involving explicit choices about whether logic will have just or unjust consequences for users. While legal approaches evaluate user interactions with algorithms as contractual exchanges that can be broken when they violate certain criteria, technical approaches presuppose that these violations can always be located in algorithm logic. This neglects the paramediation of phenomena, which cannot always be evaluated according to a definitive truth value, or explained in terms of algorithm logic. From the perspective of paramediation, this approach reaches the same limits as to the task to identify and censor deepfakes produced by I2I.

As the algorithm scholar Tyler Reigeluth has argued, if we evaluate recommendations according to whether they are good or bad for users, we neglect how these systems participate in shaping the way that users perceive themselves (Reigeluth, 2017). This means that the implications of recommendations can't only be understood in terms of their logic and appropriateness, and require an attention to what they evoke in our perception of ourselves. Reigeluth's critique poses a challenge for legal approaches to social media algorithm regulation, which target specific algorithmic processes of discrimination and misuses of data. It complicates technical approaches which attempt to quantify and optimize the quality of recommendations without acknowledging changes in human subjectivity that are not reflected in these heuristics. And it challenges criticisms that attempt to identify the harms of recommendations in algorithm logic, which are undetectable without a regard for the perceptions and experiences that inform human subjectivity.

The tendency to reduce the harms of algorithms to the consequences of their logic, or to deceptive interface design practices, can also be challenged by shifting attention to a phenomenology of paramediation. Rather than classifying particular consequences of algorithm logic as harmful, and attempting to work backward to locate their origin in logic or particular design components, a phenomenological approach addresses how algorithmic techniques and distributed percepts are configured – sometimes incidentally and automatically, sometimes intentionally by design, and usually both – to influence human perception and subjectivity. For a theory of paramediation, these configurations are at play irrespective of our efforts to make algorithmic systems more fair, transparent, or harmless.

Rather than demanding that users be protected from misuses of data, opaque algorithms, and discriminatory recommender systems, paramediation acknowledges that conditions of unpredictable harm and opacity are not possible to be definitively expunged from algorithm operations. This is not to say that the harms of algorithms are inescapable, or that algorithms are inherently harmful; only that there is a limit to how transparent, equitable, and benign we can make algorithmic systems by optimizing content quality heuristics to improve recommender systems. A more robust approach to confronting the social consequences of algorithms would need to address how these systems configure phenomena over time, and in relation to user perception, which involves a greater indeterminacy and relationship to subjectivity. To better ensure that these systems do not cause harm, we need to account for the conditions of perception they tend to evoke – by design and by accident.

The goal to improve algorithmic recommendations and optimize their appropriateness loses sight of paramediation, when it aims to design algorithms that contribute to more positive user perceptions of them. This may only shift the harms of paramediation elsewhere, ensuring that it can manage behavior more effectively, while appearing less suspect to users. More appropriate recommendations may minimize grievances about inappropriate or biased recommendations, while effectively obscuring less salient algorithmic configurations of perception, association, and subjectivity. In contrast, an attention to paramediation remains concerned with the conditions of perceptibility that algorithmic systems enable – especially when they are designed to appear less harmful. It motivates us to ask not whether particular recommendations are appropriate, but the effects that recommendations can be designed to have.

3.2.3 Palantir Gotham

The goal to improve recommender systems raises an important aspect of algorithm design: the ability to design algorithmic systems such that they alleviate concerns about their possible harms. This ability is significant for the design of systems that are notorious for their social harms, such as algorithmic systems used in law enforcement. In this section I examine Palantir Gotham, a crime data analysis platform used for conducting criminal investigations. I choose Gotham for analysis because, in contrast to other algorithmic systems used in law enforcement operations, the use of Gotham has not been significantly impeded by critiques of algorithmic bias and opacity. This suggests that it may be more effective at resisting these critiques, and that another approach to evaluating its effects is in order.

Like TikTok, the specific operations of Gotham's have not been disclosed. But even beyond this, the software itself is proprietary; only paying law enforcement institutions have access to using it. For this reason, Palantir Gotham is a useful case for identifying what can, in fact, be known about black boxed algorithmic systems that cannot be reverse engineered. In previous work, I demonstrated that much can be learned about these systems by attending to their public documentation and design specifications (Polack, 2020). These design documents can be seen as documenting the design and motivations for paramediation, enabling us to grasp the operations and social consequences of algorithmic systems by attending to how they are designed to be used and perceived, when information about their source code is not accessible.

Another significant implication of Palantir Gotham for this analysis is that it does not determine the outcomes of crime analysis without human input: people and algorithms interact to derive information about criminal investigations. This frustrates critiques of algorithmic systems that try to hold algorithm logic or human biases responsible for their harms – the effects of Palantir Gotham reside somewhere in between both, irreducible to either. Another approach is in order to help us to grasp this interval between the agency of humans and algorithms, between the computed norms of crime data and the paramedia it produces.

3.2.3.1 How Palantir Gotham Works

Palantir Gotham represents relationships among people, places, objects, events, organizations, and crimes – all called "entities" – as a network diagram. Each entity has its own "profile" of information which can be added manually and edited. Crime analysts identify relationships between entities by adding information about them and interacting with the network diagram that displays their relationships, namely by adjusting parameters that generate the diagram. The generated diagram could be considered an instance of paramedia: while it reflects the 'ground truth' of crime data, it simultaneously adds something new that deviates from this basis. This newness is not simply a modification of a statistical norm like I2I, nor of a causal norm like exposure on TikTok, but of relational norm between data elements that is thoroughly customizable. This is to say that crime analysts can edit the norm which defines criminal associations in Gotham. More than a flexible norm, this is an interactive one.

Initially, Palantir Gotham accomplishes this with "record linkage" techniques that automatically identify relationships between diverse data elements. These algorithms range from more rudimentary "approximate string matching" algorithms to more sophisticated machine learning algorithms. Approximate string matching is used to automatically identify whether two elements of text (strings) refer to the same term, even if they have different spellings or misspellings. This enables large databases to be simplified and organized, by eliminating redundant data entries and identifying relationships between multiple uses of the same or similar term.

Other record linkage techniques extend the logic of approximate string matching further, by identifying approximate matches between whole data entries and documents. A significant use of this approach in crime analytics is for "crime-suspect correlation," which attempts to automatically associate data about crimes with data about people. Crime-suspect correlation is an area of research that concerns how patterns in data can be used to identify perpetrators of crimes – especially repeated crimes – that human analysts would normally miss on their own. For example, a crime-suspect correlation method may identify the repeat occurrence of a crime with a particular weapon and at a particular time of the day, leading the algorithm to estimate a relationship between these crimes, this weapon, and a criminal suspect who is known to use it.

Altogether, Palantir Gotham could be viewed as a crime-suspect correlation platform in its entirety: the software is designed to use automated algorithmic logic and human interactions to link crime data with criminal profiles (or with profiles of people who are not suspected of crimes, notwithstanding). But unlike traditional crime-suspect correlation techniques that are implemented entirely by algorithm developers, Palantir Gotham enables crime analysts to participate in the process of specifying the parameters and constraints of the correlations. The appeal of Gotham to crime analysts is that it makes crime-suspect correlation more flexible, opening it up to more manual decisions and interactions from human users. Whereas automated crime-suspect correlation infers a norm from data, Palantir Gotham enables this norm to be customized after the fact.

Record linkage can occur either by combining multiple data entries into one record (e.g., combining multiple documents because they refer to a single crime), or by linking data entries while retaining their distinction (e.g., linking a crime document with a vehicle document). The rules for these combinations and distinctions are determined in the first instance by Gotham's designers: the platform enables data entries to be combined into single profiles about people, crimes, organizations, locations, or vehicles. These "entities" compromise the fundamental ontology of the Gotham platform, determining the kinds of entities and relationships that can exist during crime analysis. While analysts can edit the information on any entity's profile, this fundamental ontology about the entities that can exists remains static.

When databases are loaded into Gotham, record linkage automatically organizes data elements so that they reflect individual entities (the Palantir Foundry platform can be used to customize data linkage more thoroughly). Paramediation is at work here to make these distributed data elements perceptible as relational objects: while these relations are derived from real data, it cannot be said that they actually exist there. Nonetheless, Gotham's data visualization engine renders the entities to the screen as individual icons, lists, and networks that indicate statistics and inferred norms about the entities' characteristics. Entities with more relationships appear closer together, such that entities appear to form groups and proximity.

At this point, analysts can add or delete links manually to change the shape of the network, or they can customize parameters that define the network links overall, like the time period of data should be used to calculate the links. One of the most advertised features of the

Gotham platform for creating these links is the Search-Around tool. Analysts can select a specific entity and perform a Search-Around according to certain criteria, which retrieves similar data and displays it around the selected node. In this way, analysts can 'grow' the graph outward by selecting entities and performing Search-Arounds.

3.2.3.2 Palantir Gotham and Paramediation

The ultimate purpose of Palantir Gotham is to enable the creation of possible crime-suspect relations and to present them to other law enforcement agencies or analysts, such as through PowerPoint presentations. This is to say that Gotham does not operationalize the connections it is used to create, by automating them to have a direct effect on crime suspects themselves. In fact, this is rarely ever the case: crime analysis systems – from crime prediction models to criminal recidivism predictors – are concerned principally with informing how crimes and criminals should be perceived by judicial and law enforcement agencies. This means that the first and last subjects of crime analysis algorithms are the people who enforce the law, who then leverage these algorithms to make other subjects appear a certain way. By configuring the phenomena that law enforcement agencies perceive, and enabling them to configure these phenomena in turn, Palantir enables them to assign criminality to people, while also minding their public perception.

From the perspective of paramediation, what is most significant about Palantir Gotham's ontology is not that it misrepresents the truth, but how it enables given events and circumstances to be perceived. In order for this to happen, objects, subjects, and events must first be simplified as entities in an ontology. What is key about this initial step is that it does not simply define what entities are able to exist, but also their *possible interactions* and namely *how their relationships unfold dynamically*. These possible relationships do not become apparent until entities are populated with data from databases, and then linked with one another according to this data. Interactions between the entities thus appear to emerge organically without manual intervention.

The original, base ontology, which specifies which types of entities can exist, becomes an elaborated ontology, which specifies entities that exist in the world and their interrelations.

From one perspective, this transition from a base ontology to an elaborated ontology can be viewed as the result of algorithm logic, which specifies an ontology for which entities can exist and how their data should be linked. Accordingly, if the elaborated ontology reflects entities that are not thought to exist (they are not accepted as *given*), then the logic of the algorithm can be changed to correspond to the given entities more accurately. But this perspective overlooks how flexible the production of this ontology is in the first place; it is never constrained absolutely by particular algorithm logic or code. Further, the purpose of this flexibility is not to ensure that analysts can iteratively develop a model that approximates reality more accurately, but to enable analysts to iteratively develop a model that suggests possible perpetrators of crimes, possible explanations of crime events, and even *possible criminal suspects* – people who are not even suspected of committing a crime, but estimated to be *possibly suspectable*.

The ultimate purpose of paramediation in Palantir Gotham is to arrange how crimes and their circumstances appear – to compose the semblance of a crime narrative. The theoretical significance of paramediation is that algorithms are carefully designed to participate in this process of arranging appearances, by coordinating their operations with the basis of collected data and recasting them through the semiotics of crime investigation. The function of paramediation is to ensure that algorithms exist seamlessly among other crime investigation practices that law enforcement agencies normally depend on, fitting in to their semiotic environment.

Crucially, this is not achieved simply by representing people as criminals on an interface display, as if an interface or data visualization is sufficient to articulate people as criminals. While it may sometimes concern interface design, a theory of paramediation goes beyond it to address how algorithmic procedures enable certain phenomena to appear and unfold dynamically in a way that is parametrically consistent with other phenomena and human practices.

Establishing this consistency is not trivial, and depends on the meticulous design of paramediation. In the case of Palantir Gotham, paramediation is designed to balance the coherence of criminal associations with their dynamic unfolding. On the one hand, if criminal associations produced by Palantir Gotham simply reproduce obvious associations in data like matching strings, no robust narrative of crime will emerge – it will be too normal. Therefore analysts are able to expand network diagrams until their interrelations between entities appear more robust. On the other hand, if criminal associations produced by Gotham are too entropic, perhaps by indicating arbitrary connections between data of all kinds, no realistically coherent narrative of crime will emerge – it will be too abnormal. Therefore analysts can filter and trim down network diagrams until they appear more refined. Enabling analysts to increase and decrease the complexity of associations, Gotham balances between normal and abnormal relations to arrive at a paranormal diagram.

Human interaction is a significant aspect of Gotham's consequences that is irreducible to algorithm logic. Ultimately Gotham is not a visualization of record linkages or crime-suspect correlations, but a visualization of the process of designing these very associations. Gotham enables analysts to participate in what designers of record linkage or crime-suspect correlation algorithms do already; only it makes this process more fundamentally flexible. While the flexibility of Gotham's operations may raise concerns about it supporting the subjective whims and biases of its users, from the perspective of paramediation the more pressing issue is how Gotham exhibits flexibility at the level of algorithm design with the goal to aid the process of crime investigation. Decisions about how to link data by algorithm developers that would normally appear arbitrary now appear to be informed by data. Network diagrams appear to unfold dynamically over time to appear as a process of discovery and revelation, enabling a flexible paramediation that is not entirely pre-coded into algorithm logic.

As discussed, algorithmic systems designed to predict the locations of crimes or people predisposed to committing them have been critiqued for their racial biases, lack of transparency, and inaccuracies, motivating the LAPD to distinguish between racist "discriminatory" algorithms and impartial "discriminating" algorithms (Bratton, 2018). This distinction, through attempting to improve the public reception and experience of algorithmic systems for law enforcement, effectively obscures the power structures that these systems perpetuate irrespective of their logic. This includes the way that Palantir Gotham substantiates particular conditions of criminality, and the perception of something as criminal altogether. Paramediation acknowledges that this is not merely the reification of an ontology instantiated in algorithm logic; it is the dynamic, interactive, and strategically balanced arrangement of phenomena, coordinated with other given phenomena and social practices.

While an epistemological analysis of algorithm ontologies may hold that we can prevent the social harms of algorithmic systems like Gotham by making their ontologies more objective, this neglects the function of platforms like Palantir Gotham to dynamically configure perceptions of crimes and produce crime narratives. An epistemological critique of algorithms can only tend toward the expansion and refinement of techniques for producing ontologies that substantiate these narratives.

Investigating Palantir Gotham from the perspective of paramediation demonstrates how algorithmic systems depend on configuring phenomena and conditions of perceptibility in order to have an effect on the world. I argue that systems like Gotham do this irrespective of whether they have graphical user interfaces. Insofar as these systems are designed to influence the outcome of criminal investigations and decisions about criminality, they must be concerned with how crimes, and human behaviors more broadly, become perceptible. Even if these systems calculate criminality automatically, they still participate in configuring phenomena to evoke perceptions of particular actions and subjects as criminal.

3.2.4 Dwarf Fortress

Videogame research readily informs theories of algorithms and their socio-political implications. For example, by investigating how players interact with procedures in game environments, researchers can observe conditions of human-computer interaction that motivate particular human behaviors and perceptions (for example Yanez-Gomez et al., 2017). Moreover, the rise of gamification – or the design of an otherwise banal process as a videogame – has motivated scholars, artists, and activists to address the function of videogame conventions in exploitative labor practices (for example Vasudevan & Chan, 2022). Altogether, videogames are types of simulations, which have been a subject of sociological analysis and political thought for some time.

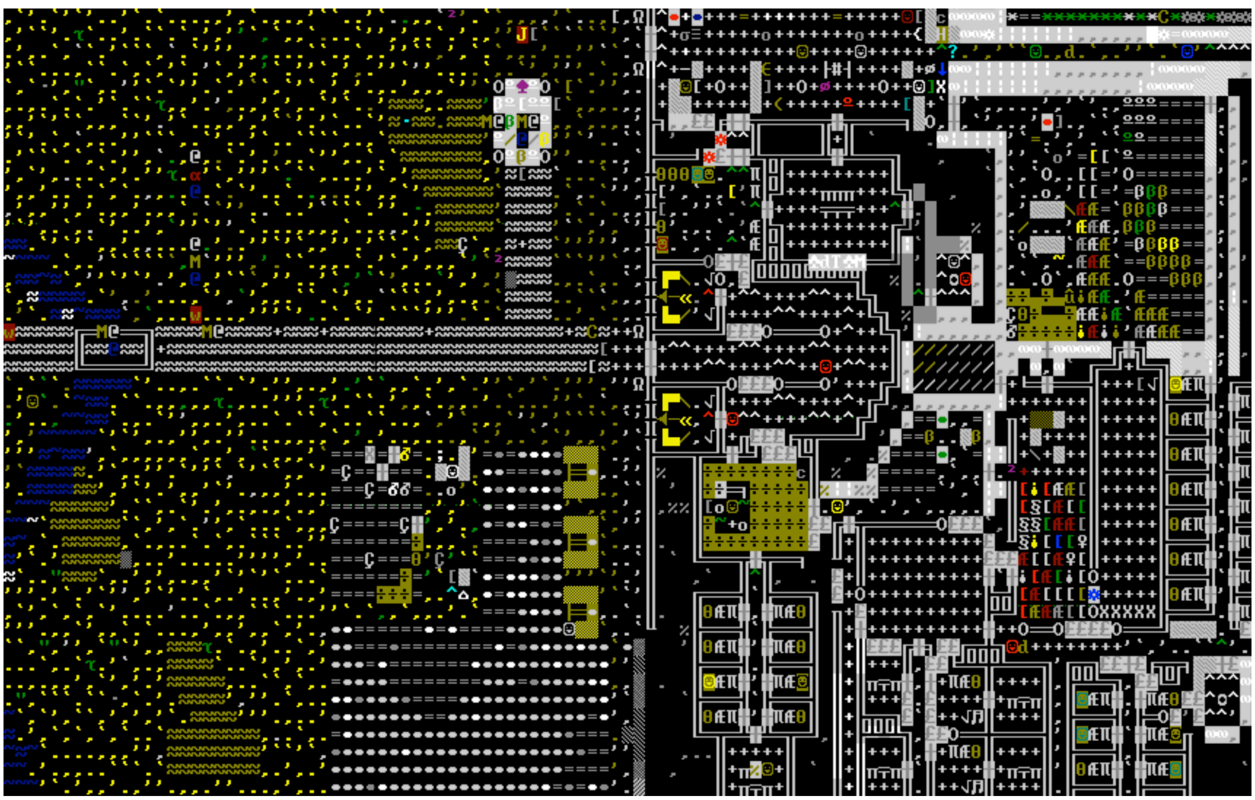

Figure 3.1: *Dwarf Fortress* by Bay 12 Games

The videogame that I analyze in this section, Dwarf Fortress, does not raise social concerns like the other case studies do – it does not facilitate a post-truth regime, exploit human behavior for profit, or participate in law enforcement operations. However, it is an important case for illustrating paramediation because of how effectively it uses algorithmic procedures to produce perceptible phenomena. In doing so, it bears important implications for algorithms in other contexts: indeed, Dwarf Frotress is concerned with (1) producing believable phenomena like image-to-image translation, (2) keeping users engaged like TikTok, and (3) producing crime narratives like Palantir Gotham – albeit with very different means and ends. Altogether the design of Dwarf Fortress (and videogames in general) motivates us to think broadly, openly, and proactively about the possible ways that algorithms can be designed to configure perception.

3.2.4.1 How Dwarf Fortress Works

Dwarf Fortress is a game where the player commands a knot of several dwarves, instructing them to perform various activities that enable them to survive in a harsh environment. The game is known for its complexity, both at the level of the game logic and the vast array of strategies that players must learn and depend on to survive. This includes knowing how to build a fortress with rooms that serve particular functions, implement sustainable food and energy sources, create an effective distribution of labor and healthy social relations among dwarves, defend against various invading forces like zombies and demons, and so on. As the game has been in development since 2002, with its two developers Tarn and Zach Adams constantly adding new features, the game's level of complexity is almost incomprehensibly large, which motivates players to keep playing the game, and to tell elaborate, illustrated stories about scenarios that their interactions with the game environment generate.

The popularity of Dwarf Fortress is all the more interesting since its graphics are extremely rudimentary, depending only on colored text characters to represent everything in the game environment – taking after the groundbreaking 1980 game Rogue. What Dwarf Fortress lacks in "graphical intensity" it more than makes up for with "process intensity," or the use of computational procedures to evoke interesting gameplay phenomena (Crawford, 1987). Following Rogue and other related classics like Nethack, Dwarf Fortress uses a wide variety of algorithmic rules to generate complex environments and dwarf personalities procedurally.

When the player starts "Fortress Mode," the game procedurally generates a vast history and geographic environment, in which the player can choose a specific location to settle. The world map is represented as a grid of colored text characters that evokes the semblance of geographical features. Players choose locations depending on the characteristics of the geographical environment; a forested and mountainous location with access to a water source is ideal for collecting resources and building an underground fortress. When the player selects a location, they see a grid of colored characters that represent various objects in the game – quotation marks for grass, the letter "c" for cats, simple smiley faces for dwarves, and other various symbols for different species of wildlife and geographical features. The symbols move autonomously according to various rules specific to each, which govern how they move and interact with other entities in the environment.

If the player does nothing, the dwarves will autonomously begin to participate in various activities, like chopping wood or looking for berries, depending on their mood and interests. The player can use various keys on the keyboard to assign dwarves to specific tasks, like mining rocks to carve out an underground cavern for a fortress. They can also inspect dwarves to learn more about them, such as their names, personal histories, and predilections. There is no specific object of the game, but the environment is generally hostile, and the game ends when all dwarves have died. Therefore, players aim to extract resources, construct elaborate defense systems, and develop a functional dwarven society that staves off boredom, crime, and insanity. This requires learning how to defend against various hostile creatures, and carefully attending to dwarves' needs, according to their each of their individual dispositions, such as building rooms with specific amenities, bars, and halls for meetings.

In a different game mode, "Adventure Mode," the player controls a single character, traveling between villages, completing quests, and learning about the procedurally generated environment and its inhabitants. As the player combats a network of villains and their incursions into villages, the game updates dynamically to reflect the player's activity in a way that retains the coherence of the narrative. For example, if the player kills the right-hand henchman of an evil necromancer, the necromancer may seek out someone elsewhere to do their bidding, or resort to depending on the next line of command. Game events and village rumors are designed to reflect this, so that the entire game environment evokes a sense of narrative consistency and progression.

The specification of rules and norms in the Dwarf Fortress environment is implemented with meticulous attention to fantasy lore, real-life processes, conditions in the game's logic, and player feedback, contributing to its overall unpredictability, and to its realism. While Dwarf Fortress is a fantasy game with no claims to being objective, the details, coincidences, and coherences of its procedurally generated environment attract players for their evocative resemblance to real-life processes and human stories. This virtual realism is then augmented by players who record their gameplay, narrate it, and produce illustrations to accompany the narrative.

When Dwarf Fortress runs, it loads a database of text files that each specify an entity and its various attributes, such as whether the entity can fly or is hostile to dwarves. These entities and attributes were designed modularly so that new entities, attributes, or dynamic interactions could be implemented without needing to reimplement or compromise other aspects of the game. Because each entity's behavior is determined by its own attributes and the attributes of other entities that it encounters in the environment, editing the game code to change any of these entity's attributes will change its behavior in the game, and other entities will respond to this behavior accordingly. These modular interactions between entities, which depend on simple algorithmic rules to generate complex chains of events, drive Dwarf Fortress's procedural environment and narratives.

Due to the game's modularity and complexity, the implementation of new features by the game's developers is likely to cause unexpected consequences. Tarn describes a circumstance where a game update unexpectedly resulted in widespread cat deaths.[6] It turned out that the cat self-cleaning behavior lead them to have alcohol poisoning as a result getting too much alcohol in their eyes that stuck to their fur when they entered bars that dwarves had spilled beer in. While some of these emergent consequences could be seen to 'break' the game by making it too difficult to play, more often they simply cause funny game dynamics or motivate more sophisticated game updates. In this case, Dwarf Fortress was patched to implement eyelids for creatures, such that alcohol would not be so easily absorbed through their eyes.

3.2.4.2 Dwarf Fortress and Paramediation

Like a scientific simulation, Dwarf Fortress begins from fundamental premises, laws, and principles that are widely accepted as given, and arranges them in a way such that they evoke phenomena that can also be deemed as given. But where a scientific simulation is concerned with the precise correlation of these rules and phenomena with the given, games are much less so – they tolerate a much wider degree of paranormality. This means that they are not even concerned if their a priori norms reflect givens; to the contrary it is sufficient if their emergent paramedia do in some roughly qualitative or aesthetic way.

This procedural paramediation of realism is evidenced by other videogames as well. Significantly, it does not necessarily depend on the production of realistic graphics, and can instead leverage techniques for designing realistic procedures and mechanisms of interaction. The procedural film artist Alex Anikina discusses how these techniques are integral to the production of videogame non-player characters, or NPCs (Anikina, 2020). By updating an NPCs

[6] https://www.youtube.com/watch?v=VAhHkJQ3KgY

dialogue and actions to reflect circumstances that happen in the game, game designers can make NPCs appear as if they perceive, think, and know about their environments. This is a critical aspect of Dwarf Fortress, where an environment of non-player characters and non-player objects evokes realism by moving and acting according to other information in the environment. Process intensity takes precedence to graphical intensity (Crawford, 1987) and still evokes a sense of realism.

If features are added to Dwarf Fortress in a *bottom-up* fashion through adding new entities and attributes modularly, the mediated effects of these features are coordinated *transversally*, to ensure that they do not compromise the intended functionality of others. Further, narrated gameplay of Dwarf Fortress makes sense of the procedurally mediated environment in a *top-down* way, interpreting the environment as if it were a real one, and not generated by enumerable interactions between algorithmic rules. The design and perceptibility of paramediation is evident on all these levels, which enables it to be configured to different ends with different effects.

In this way, Dwarf Fortress depends on similar principles of procedural design as Palantir Gotham. Both systems provide an ontological framework of entities through which more dynamic and generative actions can *play out*. Both systems use modular rules of association to generate complexity in a bottom-up way, and depend on the transversal management of associations to ensure that they do not lead to contradictory or redundant conclusions. When the algorithmic rules are optimized to balance coherence and spontaneity, a player or analyst can interact with the system to perceive a narrative from the top down. Crucially, from the perspective of paramediation, this top-down perception is not entirely removed from its bottom-up configuration at the level of algorithm logic, but is rather an essential aspect of its design and functioning.

This is not to claim that Palantir Gotham produces fictions about the given in the exact same way that a videogame does. To the contrary, whereas Palantir Gotham is designed to interoperate with the given in real time, sourcing information about entities from data collected in the world, and organizing it in a way that evokes a semblance of real events, Dwarf Fortress evokes pseudo-realistic processes without attempting to correspond to them in real time. Both algorithmic systems generate crime narratives, in fact, but Palantir Gotham uses data to license investigations into violations of the law, while Dwarf Fortress is driven by a fictional ontology of crimes, respective punishments, and randomized behaviors. From the perspective of paramediation, the distinction between the systems is how algorithmic processes and phenomena are designed to appear consistent with other given, non-algorithmic ones, like legal procedure.

Where this correspondence is not absolute, the systems depend on human perceptions to fill the gaps. While Dwarf Fortress depends on players to perceive its text-based interface as a living and breathing environment, and to supplement these sensations with narratives and illustrations that evoke credibility, Palantir Gotham depends on analysts to perceive its network diagram as an adequate expression of people and their interactions, and to edit it whenever this is not adequately invoked. Note that this is the same dependency on human perception that is integral to designing effective image-to-image techniques and influencing human behavior on TikTok. These algorithms are not simply designed to represent the truth deceptively or control human behavior, but to evoke phenomena with an attention to human perception. Even where this algorithmic logic depends on constructing an ontology at the level of code, this digital ontology cannot be effective without attending to phenomenology.

Dwarf Fortress's code is open source and its many complex behaviors are well-documented. And yet, we still cannot anticipate all the scenarios and imaginaries that will emerge from its use. This is a critical aspect of paramediation: the effects of the system cannot be reduced to algorithm logic or ontology, because this logic is designed precisely to *play out* in a way that manifests phenomena to human perception, typically in a way that is coordinated with real-life processes – aesthetically in this case, but statistically in others. Dwarf Fortress does not raise critical ethical issues (beyond, say, its caricature of social relations, crime, and carceral logic), but it remains a significant case for illustrating how paramediation is designed at the edges of perception and imagination.

3.3 Paramediation: Functions and Paraesthetics

In each of the four cases, algorithms are designed to produce perceptible phenomena and coordinate them with other, non-algorithmic processes. In each case, algorithms are used or acknowledged as an object of design, which must be configured to solve particular tasks. These tasks, in turn, give rise to specific challenges or hazards – algorithm design fails to achieve the desired paramediation – which designers employ various strategies to solve. These tasks and their apparent failures, which manifest in the aesthetics of the *algorithmic paranormal*, can tell us about the functions of paramediation, that is, what paramediation is used for. Identifying these tasks and their commonalities may help to grasp what exactly paramediation does beyond particular uses cases, toward a more generalizable theory. In this section, I identify some of these tasks and their paraesthetics that are evidenced by the case studies, to raise some summarizing aspects about paramediation before proceeding to interviews.

Function of paramediation: various image-to-image translation tasks, generally to evoke realism
Strategy: exploit the low resolution of images to deceive human perception
Paraesthetics:

Artifacts (or Data Orbs)

Image-to-image translation abstracts structural and aesthetic features from multiple images and reproduces them according to certain constraints. The "composite" images that result are recombinations of source material, which are arranged in a way that should evoke a semblance to them. The process of recombination often results in *artifacts*, or phenomena that appear visible distinct from others, or altogether nonsensical. Like the photographic artifacts that were thought to reveal phantasmic "orbs," these algorithmic artifacts attract attention for introducing something to an image that is not already accepted as given. Artifacts cannot be said to exist until code is executed and they are identified by human perception – only then can they be traced back to code and located in a particular configuration of algorithm logic. Thereafter, these configurations of logic can be correlated with artifacts, so that logic can be redesigned to mitigate them. In image-to-image translation, paramediation is designed to mitigate the appearance and perception of artifacts.

Empty Signifiers (or Cursed Images)

While artifacts are the interrupted aesthetics of composite images, empty signifiers are interrupted semiotics. Empty signifiers occur when composite images evoke the appearance of a sign, language, code, or symbol which is not properly organized to convey a meaning. For example, a composite image may produce letters that evoke the appearance of latin characters, but which are not organized in a way to have any linguistic meaning. Such characters are called "cursed" in internet parlance: they evince an uncanny feeling, where recognition breaks down. The presence of empty signifiers may suggest that an image has been manipulated, so paramediation is effective at I2I when it is designed to avoid the production of empty signifiers.

Repetition (or Overfitting)

GANs are designed to solve image-to-image translation tasks without simply duplicating input images. Image duplication is caused when a model converges on a solution that *overfits* to training data (also called mode collapse). The resulting composite images evoke a repetitive appearance that lacks variability, which can reveal that they have been synthesized by an algorithm (or merely copy an existing image). *Artifacts* may also be related to *repetition*, as they may be indiscernible until they are seen to repeat many times, and thus become identifiable as such. Paramediation that aims for realism is implemented to avoid such repetition.

Latent Signifiers (or Latent Space Cryptids)

All image-to-image tasks involve the synthesis of a composite image that does not correspond to absolutely to the given. In the case of artifacts or empty signifiers, this non-correspondence becomes apparent, insofar as the coherence of the image's aesthetics or semiotics is compromised. But image synthesis may also result in composite images that, while not aesthetically or semiotically incoherent, nonetheless appear unreal. This includes figures that consistently reappear in synthetic images, such as Loab and Crungus. Loab was named as such by the algorithm artist Supercomposite, who identified the parameters to generate the appearance of a horrific woman, whereas Crungus refers to the text input provided to the DALL-E algorithm which causes it to generate a horrific gnome creature. These figures emerge through the way that training data is encoded into an algorithm's hyperparameters or latent space, which causes them to consistently reappear provided certain constraints. This consistency conveys the sense that these latent space cryptids exist in some way, if only at the level of algorithmic processing. While algorithm artists may seek to explicitly make these figures appear, they may also appear accidentally during image-to-image translation, in which case paramediation must be designed to avoid it.

Function of paramediation: maximize user engagement

Strategy: (1) measure user interactions with content and customize content accordingly

(2) present exposure as a causal relationship to incentivize engagement

Paraesthetics:

Uncanny Accuracy

Uncanny accuracy occurs when a user feels that content recommended to them is extremely relevant, but also that they did not knowingly provide information that would enable this accuracy. In other words, uncanny accuracy is a discrepancy between what a user thought they 'told' an algorithm, and the sense that an algorithm evokes about what it nonetheless seems to 'know' about them. When the algorithm appears to 'know' more than the user 'told' it, users may feel that it has compromised their privacy by surveilling their behavior, or may feel disturbed that the algorithm 'knows' more about them than they expected. Uncanny accuracy may be the result of a chance circumstance, as opposed to the result of an intentional algorithm design. It becomes perceptible through paramediation.

Shadowbanning

Shadowbanning refers to the use of algorithms to censor content by systematically limiting its exposure, without having to delete it outright. Social media platforms are often charged with shadowbanning political or racial content, such as content associated with the Black Lives Matter movement (Nicholas, 2022). Social media companies typically respond by arguing that systematic shadowbanning is not at play, and that an algorithm has simply limited the exposure of certain content automatically and inadvertently. Key here is the fact that shadowbanning is primarily reported by users who perceive it, rather than by analyses of algorithm logic – it is something evoked. This should not delegitimate concerns about shadowbanning, but encourage us to understand how paramediation manifests as perceptions of censorship and neutrality.

Algorithmic Discrimination

Whereas algorithmic bias is the disproportionate treatment of individuals by algorithm logic, algorithmic discrimination might be distinguished as the appearance of this prejudice to human subjects. For example, concerns raised about TikTok algorithms grouping people by race or attractiveness do not point to specific biases in algorithm logic, but rather to the perception of algorithmic discrimination. Or, algorithmic discrimination may appear in I2I, as an evoked bias toward certain phenomena over others, like skin color, in composite images. The fact that algorithmic discrimination may be merely perceived, and not verifiable at the level of algorithm logic, does not mean that it is less harmful. Instead, it demonstrates that discrimination by algorithms can be happen phenomenally in ways that are irreducible to algorithm logic alone. If algorithmic bias is found in logic, algorithmic discrimination is found in paramediation.

Task: evoke crime narratives substantiated by data

Strategy: (1) use a reductive representation to make arbitrary data and logic appear to evoke real events (2) implement modular entities with many attributes that evoke unfolding interactions (3) use change to evoke meaning

Paraesthetics:

Interference

Crime analysis software like Palantir Gotham proposes to represent the given (e.g., where a crime has occurred, who is a criminal) to crime analysts so that they can identify perpetrators of crimes. If algorithms or crime analysts are able to *manipulate the representation of the given*, it would compromise the platform's impartiality and claims to the truth. Nonetheless, this is precisely what software like Gotham is designed to do: to enable the manipulation of crime data in order to evoke a crime narrative. Palantir Gotham must therefore be designed so that the algorithmic and interactive organization of crime data does not appear to involve *excessive* interference by algorithms or analysts in ostensibly positivist calculations. Certainly, what qualifies as excessive interference is subjective. Nonetheless, we can acknowledge that disbanded crime analysis platforms like PredPol failed this test, by conspicuously involving an excessive amount of human and algorithmic interference with what is otherwise regarded as given. Paramediation in crime analysis is thus designed to mitigate the perceptibility of interference with the given – algorithmic and human alike.

Misrepresentation

Palantir Gotham proposes to represent the existence of entities, events, and processes. If this representation appears to obviously conflict with the given, it can be suspected of misrepresenting it. Paramediation can be designed to avoid the phenomenon of misrepresentation by using reductive, figurative, and graphical representations that appear to evoke given circumstances without representing them comprehensively. To do so, paramediation may be strategically designed to avoid the following superfluidity and disambiguation:

- **Superfluidity**. Superfluidity occurs when an abundance of information compromises its capacity to evoke accuracy and precision. For example, while it is possible to connect all entities in Gotham with one another, this would result in superfluous connections that compromise their appearing to evoke reality. Associations must therefore be filtered and pruned according to various criteria.
- **Disambiguation**. A disambiguated data element is one that is clearly associated with its provenance: it indicates when, where, how, why, and by whom it was collected – how precisely it has been *given*. Palantir Gotham depends rather on *ambiguating* these elements so that these details do not compromise the investigation of higher-level, more ambiguous patterns. This can be accomplished through data visualization or geographic mapping, which both aggregate multiple data elements from diverse times and locations into a flattened graphic. It is also achieved by representing specific associations between entities as lines. If the diverse ways that this heterogeneous data were given were made less ambiguous, it would compromise the coherence that the composite image evokes. Paramediation of crime data thus depends on the flattening, spatializing, aggregating, and blurring of specific data.

Misrepresentation is inherently related to *interference*: misrepresentation implies some form of interference by algorithms or analysts.

Function: evoke an entertaining environment that evokes aspects of reality

Strategy: (1) use a reductive representation to make arbitrary data and logic appear to evoke given events (2) implement modular entities with many attributes that evoke unfolding interactions (3) use change to evoke meaning

Paraesthetics:

(Un)predictability

While absolutely random and unpredictable game elements would prevent players from being able to learn how to interact with them (similar to *superfluity*), absolutely determined and predictable game elements would no longer be entertaining (similar to *repetition*). Dwarf Fortress balances these extremes by using procedural generation and modular design to evoke unexpected behaviors through constraints.

Disjuncture

An entity that fails to respond to given information in its environment evokes a disjuncture. For example, if a knot a dwarves is attacked by a giant rat and one fails to respond, this evokes a disjuncture between the given information that a giant rat is attacking and the appropriate response of attacking back. Disjunctures compromise the ability of algorithmic display to evoke complex phenomena, like simulated intelligence, life, and social relations. A disjuncture is similar to a glitch or a bug, but whereas a bug compromises the appropriate functioning of game logic, a disjuncture compromises the appropriate paramediation of phenomena. For this reason, some disjunctures can be entertaining, insofar as illogical or unexpected behaviors can evoke surprise or humor. The low graphical resolution of the Dwarf Fortress environment can also make disjunctures evoke other phenomena that do not exist at the level of code. For instance, a dwarf that fails to defend herself from a giant rat might evoke the idea that she is immobilized with fear, rather than evoking a technical failure to correlate information in the environment.

Miscommunication

A different kind of disjuncture occurs between a user's given input and a system's output, rather than between information within the system. If a user does something and expects a certain response, but the system returns a completely different response, they will perceive miscommunication. In human-computer interaction design (HCI), miscommunication is attributed to a lack of correspondence between a system's internal logic and a user's "mental model" of it. HCI therefore insists on the design of "feedback" mechanisms and clear design conventions that evoke possibilities and results of user actions. But the conditions of miscommunication in game design, and in paramediation, are more complex. In games like Dwarf Fortress, miscommunication can occur when the player commits a certain action with the intention of realizing a certain effect – for example, by attacking a violent dwarf to neutralize a threat – and then realizes that their action does not achieve the effect they intended.

Chapter 4: Study Design

4.1 Data Collection and Analysis

To develop and evaluate the theory of paramediation, I interviewed specialists whose work visibly concerns how algorithms manifest to human perception: artists, experience designers, community educators, and game designers. Specialists in these fields can be considered as organic algorithm phenomenologists: their work involves an analysis of how algorithms are perceived and designed to be perceived, even if they do not indicate this explicitly when describing their work. The understanding that these specialists develop of algorithms, their appearances, and their perceptions can thus help to develop a phenomenological approach to algorithm analysis, as well as to evaluate, inform, and clarify a theory of paramediation.

The interviews took place in two rounds. In the first round, participants were asked a series of questions about how they use and conceptualize algorithms in their work, as well as challenges involved in doing so (see Appendix B), to evaluate whether whether aspects of paramediation are evident in their work. Following this first round of interviews, I identified common themes and theoretical issues for further investigation (analyzed in Section 5). Then, in an attempt to illustrate these considerations, I designed an interactive game that illustrates some of these themes and issues (analyzed in Section 6). In the second round of interviews, I invited the participants to play the game and discuss their experience with it (see Appendix D).

For the first round of interviews, the object of analysis is not only how the participants perceive algorithms, but the material conditions and practical problems they consider when designing algorithms to make things perceptible. The emphasis is on *how* perceptible phenomena are configured, as opposed to *what* is perceptible in the last instance – although any discussion of one necessarily involves a consideration of the other. This part of the study could therefore be regarded as a post-phenomenology (Ihde, 2008): it insists on attending to material processes and social practices that make phenomena perceptible, as opposed to concentrating exclusively on the interpretations of a perceiving subject.

After the first round of interviews, the design of the game was informed by participant responses and the case studies from the last section, according to a research-creation methodology (Chapman & Sawchuk, 2012) applied to game design (Westecott, 2020). The purpose of the game is first to examine the process of designing paramediation in more detail, and second to analyze the capacity of algorithmic media to demonstrate considerations of paramediation. The purpose of the second round of interviews is to assess this capacity: I invited participants to play the game I produced and describe their observations according to a think-aloud protocol (Hoonhout, 2008). Following gameplay, I interviewed participants about their experiences with the game.

4.2 Participant Recruitment and Data Collection Protocol

I selected two people from each of the four specializations – art, experience design, community education, and game design – for a total of eight participants. I chose these specializations because each exhibits an attention toward human perceptions of algorithms, as well as the conditions that make algorithms perceptible. While I argue these considerations are relevant to designing algorithms in all circumstances, they are particularly salient in these four fields. Following the analysis of the study, I discussed why this may be the case.

I used snowball sampling to recruit participants, using different criteria to select participants in each setting:

1. Art – I began by recruiting artists from my personal network. I planned to find one artist whose work addresses human image synthesis, or which involves images of human likeness. Beyond this, I planned to recruit another artist has dissimilar work from the first.

2. User experience design – I began by recruiting experience designers who conduct research in "algorithmic experience design," who are authors of publications in this field (for example, Alvarado & Waern, 2018). Beyond this, I prioritized selecting experience designers whose work involves machine learning algorithms, engagement design, or data collection.
3. Community education – I began by recruiting participants from the Stop LAPD Spying Coalition.[7] I focused on this organization because I am seeking community educators who have developed pedagogical materials about algorithmic systems. I focused on "grassroots" community education as that which raises non-expert experiences and voices.
4. Game design – I planned to recruit two game designers who have dissimilar work from one another.

I invited each participant by sending them an introduction to the study by email. I provided my name, affiliation, contact information, and a brief introduction to my research project. This introduction stated that I am studying perceptions of algorithms in the recipient's own field, and included a brief summary of the interview process, which includes two hour-long interviews. I then asked whether the recipient would be interested in participating in the research, and I noted that interview audio and video would be recorded. If the recipient expressed interest, I sent a follow-up email with an information sheet that included a documentation of the study procedure, and I scheduled a Zoom meeting with them through my UCLA Zoom account for the first interview. Upon meeting with the participant, I introduced the study again, asked if they had further questions about the information sheet and study plan that I e-mailed them, and then obtained their oral consent to participate in the study, prior to beginning the interview.

[7] https://stoplapdspying.org/

4.3 First Round of Interviews

Through the first round of interviews, I aimed to answer a series of research questions (see Appendix C). When the participant joined the remote meeting, I introduced myself to them, reviewed the plan for the interview, and asked them if they had any questions or concerns. After an informal introductory conversation, I asked the participant if they would like to start the interview, if they consented to the audio and video being recorded, and if they would like their data to be anonymized. Interviews were conducted over video in order to allow participants to screenshare their work. During the interview, I took notes on a network-disabled note-taking application. After each interview, I took summarizing notes and transcribed the audio recording. I used Zoom's recording function to record audio and video and saved the file directly to an encrypted password-protected hard drive. In the case that a participant chose to be anonymized (none did), I planned to record audio only while their face was visible on the camera, and to switch to audio-video recording when they shared their screen.

Each interview in the first round lasted an hour. To design interview questions and analyze participant responses, I followed James Spradley's research method for ethnography (Spradley, 2016). This first involved designing descriptive questions, which prompted interview subjects to describe their experiences with a disciplinary practice that they are accustomed to. In this study, participants were prompted to describe their work with designing and observing algorithms; accordingly I applied Spradley's methodology to addressing this specific aspect of the participants' work. The interviews were semi-structured, following a series of prompts that I developed in advance, which are informed in part by the participant's work (see Appendix B). When participants discuss their work, I invited them to screenshare the work in question, or any documentation of its design process, while they discussed it.

To analyze the notes and data from the interviews, I began by coding participant responses to group them into thematic categories. Following Spradley, I first identified "symbols" that interview subjects used to account for their practices and perceptions. In order to clarify the significance of symbols, I asked how they are used in practice. As well, to account for material that participants screenshare, I noted visual and nonverbal cues that correspond to a participant's verbal report. This included, for example, illustrations used in the process of designing algorithms, graphics used to explain how algorithms operate, or software environments. To code participant responses, I first identified verbal and nonverbal cues that are present across multiple participant responses. Second, I identified cues that correspond or challenge any of the theoretical considerations of paramediation.

As I coded participant responses and symbols, I took notes on similarities, differences, and idiosyncrasies observed across the four specializations, and organized them into a series of themes that reflected issues raised throughout the interviews (Section 5). After coding, I used the organized data to answer my research questions (Appendix C). This required that I analyze how participant responses correspond to the theory of paramediation I developed in Chapter 3, as well as to the list of paramediation functions and paraesthetics that I identified. Participant responses would either evidence or challenge the fundamental considerations of paramediation, and they would either reflect the paramediation patterns identified in Section 3 or demonstrate new ones. This analysis thus enabled me to examine how participant responses related to the theory or inform its premises.

4.4 Designing an Interactive Game

Following the first round of interviews, I participated in the design of paramediation by producing an interactive game. By game, I mean an interactive algorithmic system that leverages conventions of human-computer interaction and digital simulation to demonstrate consequences of human input, and which elicits play. The task of designing a game served two main purposes in this study. First, it enabled me to document and analyze firsthand how paramediation is designed. While the interviews provide an account of this, engaging directly in the design of algorithms for paramediation in a game revealed additional design considerations involved.

To accomplish this, I followed a research-creation methodology (Chapman & Sawchuk, 2012). Research-creation is an approach to practice-based research or Research Through Design (Stappers & Giaccardi, 2017), where a researcher produces knowledge through the process of designing an artifact. Instead of researching how to improve the design process (e.g., how to design the best game), research-creation involves identifying functional and aesthetic resemblances between objects of analysis and a made artifact. In this study, the purpose of making a game through research-creation was first to analyze the game design process as an instance of designing paramediation, and second to investigate how algorithmic media can be designed to demonstrate properties of paramediation. The design of the game was motivated by the theory of paramediation and was informed by participant responses. By developing theoretical and practical considerations into a game, I analyzed resemblances between these considerations and the process of implementing algorithms for paramediation in practice.

There are several reasons for implementing a game, as opposed to another type of algorithmic system:

- Games, when they are played, can simulate the process of configuring or designing algorithms.
- Games involve the composition of many perceptible phenomena, such as visual images, spatial environments, audio, and interactive feedback. Games exhibit paramediation conspicuously.
- Games permit a wide variety of human-computer interactions, and involve the design of a wide variety of algorithms. Moreover, game design is rich with diverse conventions for making interactions between humans and algorithms perceptible.
- Games enable the composition of fiction via interactions, which enables a game designer to attribute real-life implications to procedures and percepts.

- Games involve elements of certainty, uncertainty, spontaneity, perspective, progress and more that are relevant to considerations of paramediation, and moreover are designed to be salient to human perception.

When research-creation is applied to game design (Westecott, 2020), the process of designing a game is as important as the game's final appearance. Westecott's approach to research-creation for games that takes the concept of game sketching as its point of departure. Game sketching is a process of envisioning desirable or practical features of a game, as well as their potential reception by game players. This is to anticipate the effects of gameplay on human experience, in advance of its production. Unlike user-interface design, the complexity of game environments makes these possible effects on human experience difficult to anticipate systematically and to distill into design conventions that can be applied as general rules. Therefore, another reason for designing a game through research-creation in this study is to examine how the open-ended practice of game sketching (that is, of anticipating human perceptions of an algorithmic system) are involved in the design of paramediation.

The second function of designing a game is to investigate how considerations of paramediation might be made more intelligible. Here we move from "creation-as-research" to developing a "creative presentation of research" (Westecott, 2020). To design the game, I began by reviewing the paramediation functions and paraesthetics I identified in Chapter 3, identifying whether they correspond to data collected from the interviews. I planned to design one part of the game for each of the paramediation functions that were evidenced in the case studies and participant responses. Each game part, or scene, would simulate how a particular function of paramediation operates. To examine whether the game was capable of demonstrating, elucidating, or even challenging theoretical considerations of paramediation, I ultimately conducted a second round of interviews.

4.5 Second Round of Interviews

Once I developed game sketches and prototypes of the interactive game, I sent a follow-up email to confirm whether the participants would like to continue with the interview process and playtest a game. The purpose of this second round of interviews was to explore the game's capacity to demonstrate, elucidate, or even challenge theoretical considerations of paramediation.

At the beginning of the second interview, I obtained oral consent from the participant to participate in the study as documented in the information sheet provided to them by e-mail. Interviews were once again conducted over Zoom with video in order to allow participants to screenshare. I used Zoom's recording function to record audio and video and saved the file directly to an encrypted password-protected hard drive. At the interview, the participant was provided a link to play the game online (hosted on a DigitalOcean server and accessible via SSL encryption), and was instructed to share their screen while they played and describe their observations following a think-aloud protocol (Hoonhout, 2008). In particular, I told participants, "as you're playing, please say aloud everything that comes to mind – whether it's an observation, something that the game reminds you of, a challenge or obstacle." During gameplay, I occasionally reminded them to continue to share their thoughts.

During the interview, I took notes on a network-disabled note-taking application, recording participant gameplay and logging the times that particular events happen in the game, for later reference and correlation with their verbal report. During gameplay, I noted participant actions in the game, their reactions to the game experience, and the sequences of actions that they conducted over time. After 30 minutes of gameplay, I asked the participants a series of questions about their experiences with the game, and whether these experiences elicit any relationship to their own work (Appendix D). During this interview, I described the theory of paramediation to participants and my goals in developing the game, following a script template (Appendix E).

Following the second round of interviews, I analyzed participant gameplay and verbal reports to further develop the theory of paramediation. To accomplish this, I followed Hoonhout (2008) to divide the recording of each participant's gameplay and think-aloud into "chunks" that refer to particular events and interactions in the game, which I will then categorized so that I can compare similar chunks across participants (Hoonhout, 2008). In some instances, a chunk was assigned to multiple categories. To analyze the 30-minute interviews that follow 30 minutes of gameplay, I analyzed and coded participant responses by following the procedure used in the first round of interviews. As well, I associated participant responses with particular categories of chunked gameplay, when they refer to them.

Once all of the participants' data has been chunked and categorized, I analyzed the chunks and coded verbal reports to examine participant responses to the gameplay. I identified themes, topics, and points of feedback that were related to the gameplay experience or the theory of paramediation. I highlighted sentences that referred to these points, and organized them thematically.

5 Interview Analysis

In Chapter 3, we saw how paramediation may be involved in four applications of algorithm design, and examined some ethical implications of applying the theory to each case. But how consistent is this account with the work of people who design and evaluate these systems? Do they also consider how algorithms are designed to arrange appearances that manifest to human perception? Do they acknowledge any tensions between this design practice and its moments of failure?

The following analysis of the first round of interviews is organized into coded themes that were identified during analysis of the interview data. After reading through all of the interview data and coding participant responses according to themes, I organized these themes to identify those that were noteworthy or shared by all of the respondents. In the analysis below, for each theme, I specify which of the specializations are involved in discussing it: designers of algorithmic experiences, organizers against the harms of algorithmic law enforcement technology, artists that use algorithms, and game designers.

5.1 Design

A key object of interest for all of the respondents was design. But the relationship that the participants have to design is different in each specialization. The designers, of course, participate in design themselves – even when they are not designing algorithms themselves, they are designing how algorithms are used and perceived. Nonetheless, they do not just take the idea of design for granted; instead they may be self-reflective about the function of design in their work. Oscar Alvarado Rodríguez, an interview participant who coined the term "algorithmic experience design," elaborates the meaning of design through describing how his approach differs from traditions in human-computer interaction:

> *...the stereotypical perception of the computer science approach in this field is: I try out this interface, and I compare different alternatives. Post-design, right? So what I mean with post-design is, I design first, and then I try it out. Following the traditional classic user-centered design approach, right? But then when you come up with the other side of the field, a more participatory design approach, then you don't design, right? You first explore with people, and then you give them tools.*

Oscar describes how "design" involves multiple kinds of approaches, where designers can relate to people in different ways. He even states that in the case of a more "participatory" approach to design, "you don't design" – at least not without first consulting people and addressing their needs.

He explains this by contrasting his own participatory approach to the notion that there is an "expert designer," which is someone who thinks that

> *...I know what people need, and then I try it out to see which is which of my solutions are the best. Instead of doing that, I give the people the words and the space and the moment for them to express themselves, and they say what they would like to have. And based on that you start playing around from that, right?*

This is not to say that Oscar does not consider himself a designer at all. He is still interested in developing "an ethical design" practice, "and even a political perspective on design, and a social perspective on design." To this end, he is concerned not just with evaluating the experiences of algorithms, but asking about algorithms:

> *what are they doing? And how can we improve them to make them meaningful for our lives? And for our societies as well.*

Oscar distinguishes this practice from what he refers to as a "design perspective," which for its part presumes to be able to

> *come up with similar rules that we can follow and findings that are always there and very strong that we can implement, always right. But I wouldn't say that we should stop asking people about how they'll deal with and how they interact with this technology – after even if I design a proper algorithmic experience – how this is changing their lives, not only in the short run, but also in the long run.*

In contrast to the traditional "design perspective" Oscar attends to the long-term and inconspicuous effects of algorithms on peoples lives. For him, this is intractable to codify into general rules, which risks neglecting user needs and experiences.

For the organizers, the relationship to design is very different: instead of participating in designing algorithms themselves, they are concerned about the other people who have the authority to design them. While they do not use the word "design" frequently, it does come up. Jamie Garcia, an interview respondent who organizes to educate communities in Los Angeles about police surveillance and data analysis technology, characterizes design as something that influences the social consequences of algorithmic systems:

> *...in policing, the algorithm and its output is completely manipulated, absolutely manipulated by crime analysts – they have the ability to edit it to choose and to decide who and what – and they don't even have to use the technology to the tee... They don't. They don't use the technology the way it's supposed to be used, the technology is already flawed by design, and in the end they manipulate it anyways.*

By saying that an algorithmic system "is already flawed by design," Jamie is responding the notion of the algorithmic "feedback loop" – the idea that policing algorithms are harmful because they perpetuate biases encoded in data. Jamie argues that "that's part of the story. But that's not completely it." For her, the "feedback loop" argument fails to recognize how algorithmic techniques are designed to justify police practices from the first instance – the harms of algorithmic systems are established by design, and not just bad data. "Design" here thus refers to a kind of intentionality in determining how algorithmic systems operate, which is not confronted by reforming their data alone.

The way that the game designers discuss "design" is much broader. The game designers mention "design" as referring to a set of shared practices, or a community of activity, that contributes to the development of an algorithmic system. The game designer and interview respondent Neilson Koerner-Safrata is interested in addressing

> *kind of like this lineage of design moments that we talk about with like a very deep reverence, and machine learning is now one of them... there's all these like logics of design and singular moments that then changed the logic of how we thought about design... the algorithm just fits into a lineage of kind of like design moments and logics.*

This "lineage of design moment" is something that exists in the material world, but that people can also participate in and intervene into.

Neilson considers the relationship between game designs and modders, who are players that unofficially appropriate and modify these designs, as he did when creating his mod of the game CounterStrike, which he named Dustnet:

> *...the folk logic of bunny hopping for example and... you could say like the movement algorithm – where these things kind of were passed around different games where ideas and concepts of what created good feel or game design – came from Quake for example, bled into CounterStrike, and all these different things. So a lot of what was added (in Dustnet) was more like also these contributions of different games and like the side communities of CSGO [CounterStrike: Global Offensive] surfing, which is now like very*

well loved and documented, and how these different ideas bled into every game and with an influence, like further designs or modding to those games.

Here Neilson describes how simple algorithms like "the movement algorithm" get appropriated by players in creative ways, as in "the folk logic of bunny hopping." These algorithms and player relationships to them contribute to a sense of "what created good feel or game design," which become codified in game design conventions that are passed from or "bled" from game to game. In turn, modders participate in design with "further design or modding to those games," in reference to existing designs, or playing with their characteristics.

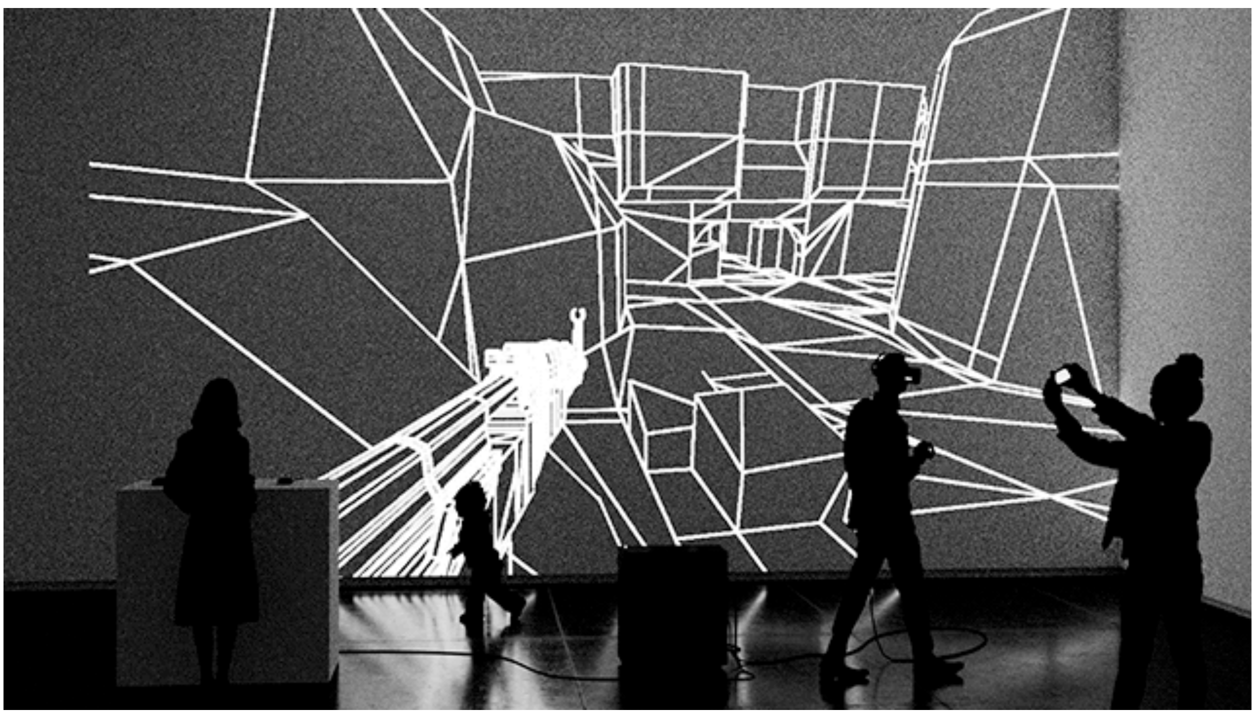

Figure 5.1: *Dustnet* by Neilson Koerner-Safrata (image courtesy of the artist)

Francis Tseng is another game designer interviewed for this study, who also has a history as a designer. Like the organizers, Francis describes how algorithmic systems are "designed" by certain people to affect others, who for their part have little control over these designs. But his specific interest in this fact is the way that people find creative loopholes in these designs or appropriate them to unexpected ends:

> *...a lot of kind of human resilience and innovation really comes up, where people find these weird ways to make do, or change things in a system that is very strongly designed to not allow those kinds of possibilities. So I guess there's an analogue there [in making simulations and games] that I really like... But it's, you know, when people – when states and stuff – design policies ... so that only a certain kind of person can benefit from something, or something like these kind of really archaic and convoluted [algorithmic] targeting methods, and, for example, social programs, or something like that – people find ways around it, or ways to evade it and stuff like that. And so this constant frustration of human ingenuity and state control, I think, is also kind of an interesting place where this comes up in the real world.*

Here Francis is describing the basis of his interest for developing simulations and games which involve procedural systems of rules: these algorithmic systems involve codified rules which operate in a way that is analogous to a hegemonic design that attempts to determine behavior. But also, these systems give way to acts of spontaneity and resistance through procedural generation and emergent behavior. "Design" here is namely something imposed on others from the outside, but also something that implies a creative practice of resistance. As a game designer, for instance, Francis experiments with these designs to illustrate and investigate how other types of behaviors can emerge.

In all cases, reference to "design" is significant because it suggests that algorithms are an object of design. While this might seem obvious, it is a fact that tends to be elided by focusing on the autonomous nature of algorithmic systems and the incidental consequences of their biases. Whereas Oscar participates in design directly as a designer, as an organizer Jamie indicates this participation is restricted to a few who impose their designs on others. Meanwhile, the game designers point to the communities involved in creating these designs, or responding to them in creative ways. Thus design is something that involves an intention to configure an algorithmic artifact in a particular way, with real effects for people, implying power dynamics as well as communities of shared designs and responses to design, which are contingent on particular algorithmic logics and their interpretations.

5.2 People

Another common object of interest for all of the respondents was people – and namely those who use, are affected by, or bear witness to algorithmic operations. In addition to using the word "people," the words that participants used to describe people differed for each of the specializations. Whereas the designers used the word "user," the organizers said "folks," both the artists and game designers were partial to "audience" and "community," and the game designers unsurprisingly referred to a "player."

The designers follow from traditions in human-computer interaction design to consider explicitly how algorithmic systems will appear to users. However, it is worth observing when they use the word "people" versus "user." Oscar, for example, uses the word "people" when he distinguishes his practice from traditional approaches. As said, against classic user-centered design, Oscar tries to "first explore with people, and then you give them tools." When describing his methods, he refers to his subjects as people:

> *...I'm going to do this interface I implemented, and then after I implement it, I try it out with people, usually with surveys, questionnaires, and also with quantitative approaches, that, based on the survey that I fill out, well, I try to do correlations and statistical analysis. And, well, I see a correlation here, in which if I put the button in this side of the screen, or if I put this visualization, people have a better sense of transparency, in contrast with this other visualization...*

In contrast, when describing the research of human behavior in human-computer interaction design, Oscar returns to focusing on the word user:

> *...we really always need users to tell us what to do. Or can we actually find specific takeaways from users that we can generalize and stop asking them? My very humble position about that debate is that we always need users – we'll always need – if we stop thinking about users, we start to talk about people. It's because it's people we're dealing with. And from a very personal stand point of view, as well, there is no reason about technology, we don't think about people first.*

Oscar is explicit that considering "users" as "people" is a way to attend to their needs, perhaps even to humanize them in the design process. By shifting from the concept of user to people, we consider how the "tools that we create and we designed" operate "for improving our lives and and to improve the lives of people as well."

Hendrik Heuer is another interview participant and designer-researcher whose work concerns whether people can understand the effects of algorithms on their lives, as well as how to design tools to aid in this understanding. Hendrik uses the word "user" when he is referring specifically to his research practice, and the word "people" when he is describing generally how people may react to algorithmic systems in various circumstances. For example:

> *I started to build like an interface for this algorithm to try to have users and they said, 'Well, it's not really working. I don't know what it means.'*

Hendrik also distinguishes between "expert users," "end users," "primary, secondary, tertiary" users who all relate to algorithmic systems in different ways. Primary users are people who use an algorithmic system in question, where secondary and tertiary users are third parties who may use a system simply for other reasons, like evaluating it in audits:

...the primary users might not be the people doing the audits. But maybe journalists will... I think auditing by third parties or secondary users, like journalists is the way to go.

Here the category of "user" is used to characterize a person's position with respect to an algorithmic system, which also entails particular kinds of expertise, intentionality, and what Hendrik calls "user beliefs," or "what's psychologically relevant to people."

Namely, Hendrik is designing tools that help people to audit algorithms, or evaluate their logic and consequences. He is sensitive to the agency that each of these user groups can have in interacting with algorithmic systems. Journalists, for example,

basically bring the insights to the end users. But yeah, I don't see, like, my parents actually doing an audit of any algorithm that they experienced. Although, it would be fun, but probably not.

Hendrik reflects on the behaviors of these lay users (not his words) often, which is where he tends to use the word "people." For instance:

And the people on Facebook don't even know that there is an algorithm ... we kind of see that people have sort of an awareness that there's an algorithm on, on TikTok, on YouTube ... most people don't know that these algorithms even exist ... it's just like, not so exciting to tell people about things that don't work. ... the primary users might not be the people doing the audits.

Reference to "people" also plays a key role in Hendrik's formulation of what he calls "the explanatory gap" in developing explanations of algorithms:

a gap between what's technically possible or easy, and what's psychologically relevant for people. And the idea of the user beliefs... [which involve] stuff that people make up, you know, like, it's just like people's perception, like, 'Maybe there are psychological *experts that like actually target me.' And we can use that, because it's kind of intuitive to people. And it relates to how they already perceive the world. And we can use that to make explanations. And then it's basically for this gap, right to fill this gap.*

Hendrik's notion of an explanatory gap (see also Heuer, 2019) highlights the discrepancy between the tools that exist to explain machine learning systems – which are for their part relatively easy to produce from a technical point of view – versus the kinds of explanations that are useful to people. What kinds of useful explanations are these? In Hendrik's words, they are those that attend to what is actually "intuitive to people," based on their existing experience.

For their part, the artists and game designers refer to "people" when they are describing how their work will ultimately be perceived. Francis, for example, discusses designing a game for the book Half Earth Socialism, where players are prompted to decide the policies that will shape humanity's response to climate change:

We're just trying to help people get an understanding of like, what kind of things we need to happen for the next 100 years to be like, relatively livable.

He considers how certain design choices will influence the engagement of people that will play the game. He states that if the game mechanics are relatively simple and not too detailed,

more people are going to play it, more people will hopefully kind of see the point that you're making.

Here there is a focus on developing an "understanding" for "people."

Neilson refers to people in the same way, in his discussion of his game-simulation Crystal Palace, where machine learning algorithms control self-driving cars in a 3D environment. He notes how he struggled to illustrate the function and aesthetics of machine learning algorithms in a way that was not exclusively comprehensible to experts:

...it was like, trying to resist that of like, 'okay, am I really just making backpropagation visible to people who understand it?'

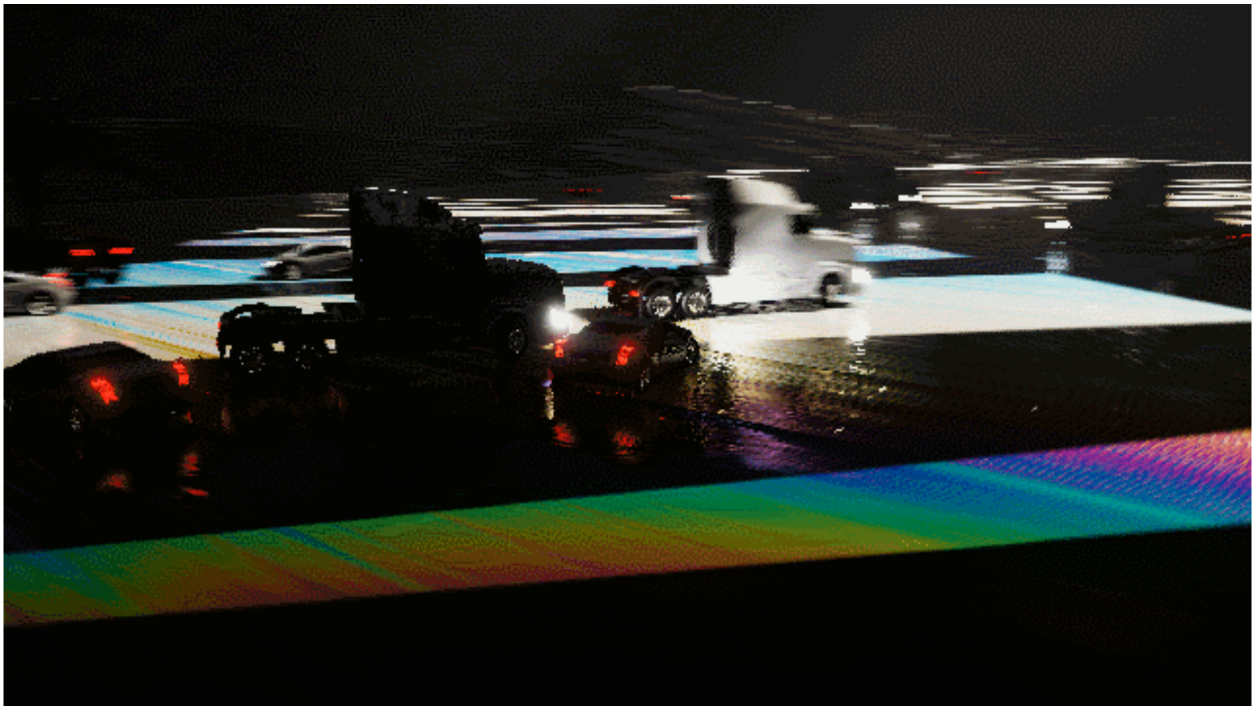

Figure 5.2: *Crystal Palace* by Neilson Koerner-Safrata (image courtesy of the artist)

Neilson is reflecting on the different types of people that can perceive his work, one being "people who understand" machine learning already. Other groups of people that Neilson refers to are "the participants" who have a history playing video games, versus those who do not:

> *...these are lived experiences for the participants a lot of the time. So they really bring all their, like, baggage to the project already in a lot of ways, so they are quite niche pieces. And even when people don't even really understand – like they haven't played CounterStrike or played Age of Empires – there is a performative aspect to them where they are multiplayer.*

Neilson is designing games so that they can be appreciated by a certain "niche" audience as well as people who "don't even really understand."

The artists and the game designers also refer to people as an "audience," and also as "communities." The artist and interview respondent Danae Tapia describes what she enjoys about her Twitter bot project, which generates tweets from her dream journal, in terms of this audience:

> *what I liked about this project... it is not me explaining everything. It's the audience who is finding meaning in the own heads.*

When I subsequently asked whether Danae designs her projects with a specific audience in mind, she answered that the audience is not specific, but that she feels

> *lucky to have access to artistic and activist spaces, where I show my work.... Also this is something I guess, nostalgic of old school internet. I like that, the digital arts in my opinion are kind of more accessible to... in terms of... to be a digital artist is way more accessible than being I don't know, a painter, or a sculptor with some crafts for which you need, I don't know, loads of infrastructure. These things are way more simple or way more accessible. And there's a community that is quite, yeah, that has fun with this.*

Danae values that her work is accessible to people that she conceptualizes as a community. This focus on accessibility and community reflects her broader interest in "hacker communities" as well:

> *... It's beautiful to see how, when people who have been knowing each other for, you know, years of talking on IRC, then they're able to meet in person. This community is very much still alive.*

The game designers also refer to communities in reference to people who creatively appropriate game mechanics to other ends. This includes Francis' interest in the "storytelling communities" that play procedurally generated games like Dwarf Fortress and then narrate their gameplay to tell stories. It also includes Neilson's discussion of gaming and modding communities. For him, one community is defined by the technical infrastructure of a game itself:

> *[multiplayer games] used to be very communal, where you kind of just selected your community server, and you lived on that thing.*

But as this technical infrastructure was replaced by algorithms that "matchmake" users with each other randomly, the community shifted elsewhere:

> *...for the last 10 years we've seen basically everything be offloaded to Discord – what used to be like happening on the server would then... just is now like a side community.*

Neilson's use of the term "side communities" distinguishes subset communities that form out of broader community of a game's players. These side communities may form out of their shared love of a game, or from their reverence of a particular game mechanic and practice, "like the side communities of CSGO surfing, which is now like very well loved and documented."

Another community that Neilson names is created through playing a more active role in a game's maintenance and development:

> *...a lot of different communities that that were all like, you know, donations paid for the server that kept it running and you confer like VIP status, or like, there's just all these entirely massive, massive communities that are completely self-made off modding, and things like that, that the Source engine allowed people to do. So like there's just a lot of time spent in those kind of communities building them.*

These communities are formed insofar as *players become creators*, modders, or designers, which for Neilson raises an important distinction. While today profit-motivated game designers tend to limit what players can design on their own terms, this wasn't always the case:

the barrier between like creator and player was... I don't want to use the word authentic to literally describe how games were played or felt, but it's just true that the barrier between like playing and creating was a little thinner and that modding and creating was encouraged back in those days and that players were more interested in how they could structure play or create play.

Policing the distinction between players and creators erects what Neilson calls a "barrier between playing and creating." In this regime, players can play with the game rules all they want – so long as they don't create something that threatens the sales of the original game experience.

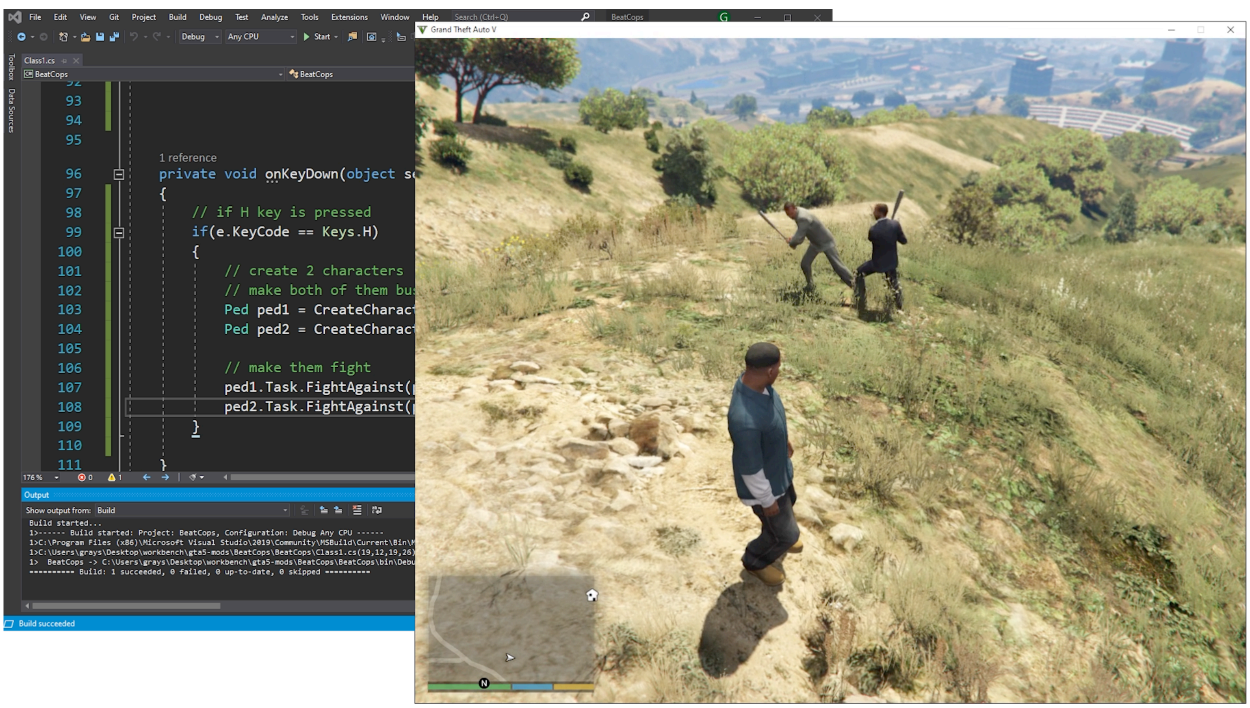

Figure 5.3: *why don't the cops fight each other* by Grayson Earle (image courtesy of the artist)

This "barrier" is also something investigated explicitly the artist and interview respondent Grayson Earle. In his project "why don't the cops fight each other," Grayson documents how his attempt to make the virtual police in Grand Theft Auto 5 fight each other revealed an unexpected discovery: the game's code explicitly forbids it. This is a fact he ultimately confirmed with a community of modders, motivating him to remark that *the modding community is this very, like interesting and pure force, because it's almost always I think, yeah, not a profit-driven endeavor, right?*

Meanwhile, we recall Francis' remarks about how "people find these weird ways to make do or change things in a system that is very strongly designed to not allow those kinds of possibilities."

On the other hand, the organizers make frequent use of the word "folks." The use of "folks" to describe people is common in activist and organizing spaces, especially where there is an emphasis on recognizing non-binary gender identities (the spelling "folx" is related to the gender-neutral spelling "latinx"; DeCarlo, 2021; Schmieder, 2022). The organizers that were interviewed also used the word "people" interchangeably, but they used "folks" when referring to a sense of everyday people, who are not specialists and may be subjected to discrimination and power asymmetries. With Hamid Khan, an organizer interviewed for this study who examines the social impact of law enforcement information systems, for example:

I'm talking about non-white communities, particularly black and indigenous and migrant, and other folks, or people who don't fit within the broader cultural or social context has been queer, trans or other folks.

For Jamie:

And part of what was so daunting about... running a campaign against predictive policing was the fact that, now we were talking about high levels of math to folks that weren't mathematicians. I wasn't a mathematician, I didn't really know how an algorithm functions.

Here folks specifies a group of people who are especially at risk of being subjected to discrimination and unfair treatment.

The word folks also recalls Neilson's reference to the "folk logic of bunny hoping," where players creatively used a game's movement algorithm to move through the game environment in a new way. Here a "folk logic" is a way of relating to algorithms that is formulated by non-experts, who may not even have access to an algorithm's logic, but find creative ways to make use of this logic nonetheless. This understanding of folk could be seen as related to research on "folk theories" of algorithms (for example Rader & Gray, 2015; Eslami et al., 2015), where the concern is how users understand algorithmic operations irrespective of their expertise.

Throughout participant responses, references to "people" – whether as users, audiences, communities, or folks – demonstrates a clear regard for the lived consequences of algorithmic systems. But more significantly, they evidence a regard for how distinct groups of people relate to these systems. Are they "primary users" who lack awareness of algorithmic systems, have they participated in playing video games before, are they "mathematicians," are they experts or folks? These considerations – about different types and groups of people – greatly inform the work of the respondents and how they conceptualize the design of algorithmic systems.

5.3 Experience

One theme that is deeply related to this notion of different groups of people is "experience." Unsurprisingly, Oscar mentions "experience" often, since it is an aspect of his theory of "algorithmic experience design." He notes that he conceptualizes "algorithmic experience" in a broad way:

> *...the idea of algorithmic experience was just to explore the field and to provide specific areas that people can start thinking about how to improve this overall thing, right? So algorithmic experience is very high level, let's say, and then you have different ways to improve the algorithmic experience, maybe in this area, in this area, in this area, in this other area, right?*

He clarifies further that broadness could be a weakness of the theory:

> *...I think that the biggest weakness of algorithmic experience was my naive approach in the beginning of thinking about algorithmic experience somehow as the user experience, right, how do we improve the user experience with algorithms? That's why I call it algorithmic experience – it's so broad as the algorithmic experience concept. And one of the weaknesses, this broadness [does not involve] a definition of the concept itself.*

This broadness has a couple implications for the theory of algorithmic experience. First, it allows the theory to account for a wide range of issues, where what qualifies as a "good" algorithmic is contingent on the object of design: "It depends on what do you design for, right?" At the same time, it leads to what Oscar calls a "trap," in reference to Paul Dourish's explanation of a "legitimacy trap," where the claim to legitimacy of human-computer interaction as a design field is reductively staked on optimizing the usability of computational systems, namely for corporate actors (Dourish, 2019):

> *...it's a validity trap or something like that. So when we say we have a good algorithmic experience, that means that technology is safe, that it's always nice, that it's the best for everybody, and that is not necessarily true, right? You can actually create a very good user experience to make people act in a certain way. But that doesn't mean that it's improving their lives, that it's improving our societies, that it's improving our way of being in the world, right. So that's another weakness of the algorithmic experience concept. So you can really have, for instance, very nice transparency feelings in an algorithmic system, and that's very good for the algorithmic experience, and then you have a very positive algorithmic experience, but the transparency could actually be false, and there are papers doing what they consider a placebo effect. So just showing something and people say, 'Oh, this is great, this is fine!' They have a lot of control. But*

> *the controls are not actually doing anything, let's say, and transparency is not actually doing anything.*

The trap of algorithmic experience design is that the effort to improve algorithmic experience can effectively detract from an awareness of algorithms and their social consequences.

Hendrik discusses this trap in the context of the COMPAS algorithm, which was designed to predict criminal recidivism:

> *I want[ed] to make that more explainable. And my supervisor addressed my thought: like, yeah, but are you sure these like things like COMPAS should exist? And I was like, taken aback. Like, yeah, but he was right! ... I mean, there's probably a lot more fundamental problems with the judicial system in the United States that need be to fixed. And algorithms are just a symptom. They're not even like the problem here, right. And to make these algorithms more explainable – I think they just shouldn't exist in a way that like, it should be a different thing.*

This "trap" of improving algorithmic experience and explainability leads Oscar to conclude that there is no way to formalize what qualifies as a good algorithmic experience in every circumstance. Instead, an algorithmic experience designer must constantly attend to the needs and experiences of people:

> *...with proper algorithmic experience, maybe we're doing something harmful for society, and we don't even realize it. And then that's, again, coming up with more inquiry, more imagination, more asking, more observation, and more studies. And yeah, I think it is a never ending work.*

Further, there is no specific design change to algorithmic systems that would put an end to this "never ending work":

Regardless of the technology behind, people have specific needs, and that's where algorithmic experience comes in, like to find out which are those needs. And you can change everything behind as many times as you want, but the needs are going to be the same, or at least [come up] over and over again, in a particular context.

This is an ongoing research practice that is not necessarily concerned with immediate design solutions:

> *...not even necessarily looking into design opportunities, but just exploring how people use technology and applications that technology are having on societies and people and so on, right. So definitely, the algorithmic experience concept is departing from those principles as well.*

At the same time, Oscar specifies some ideas about what types of algorithms should be considered in an analysis of experience, what he refers to as those that are "worthy of checking the experience":

> *... we have an algorithm that controls the mouse movement in my screen, do I need to care about it? ... No, we're not talking about that, right. We have the algorithms that are worthy of exploring their experience based on social sciences and theorists that have been dealing with some characteristics of specific algorithms that have specific implications on our societies right. ... So algorithms that are deciding what is popular, and what isn't popular, algorithms that are filtering information, ranking elements, products, information as well. Algorithms that decide what you should see and what you should not see. Yeah, I mean, those are the general characteristics of these algorithms that I consider worthy of checking the experience, but not only checking the experience, but actually worthy of reflection, worthy of considering from the design perspective, what are they doing? And how can we improve them to make them meaningful for our lives?*

Here, there is some decision that is made from the perspective of "social sciences and theorists" or "from the design perspective" about which algorithms should be studied. But in addition, researchers that attend to experience can acknowledge that the effects of algorithms can be sensed, registered and felt by people even when they are not explicitly aware of an algorithm's existence:

> *...there are things that people don't realize so much, but we do realize it from an experience perspective, which is awareness. So people, right, somehow they don't realize that there is a system selecting stuff for them. [But] I think people more, over and over, they realize it more...*

The game designers are especially sensitive to "experience," which is a way that they understand the effects of their work on people, as well as their goals for design. For Francis, for example:

> *...we're thinking about, 'Okay, what kind of experiences or stories do you want the player to experience?' And we'll work back for backwards from there.*

In other words, a regard for the experience of users informs the design of algorithmic logic through a process of 'working backwards.' These considerations of experience weigh into the development process, affecting various design decisions:

> *...with the game, the player needs to feel a sense of agency, like that they can actually influence the outcome. Otherwise, it's just a frustrating experience.*

In Neilson's work, the "lived experience" of playing games is a specific area of focus. He discusses how the experience of interacting with specific algorithms in games is something that can be controlled, taken away, and reclaimed, insofar as algorithms are commodified and proprietary:

Part of it's definitely, I think with the algorithm, is like the sovereignty over your experience with the algorithm. And I think when I do things like Dustnet, or I use, reuse game assets for art, these are all things that are on your computer or in your life, or in your lived experience, and you aren't able to actually really access these things ever again. They kind of sit in this weird space between intellectual property and like, digital archiving, which is a nightmare, right? Like you can't really have these experiences because they're so built up in like a corporate, like consumer exchange between these things. And so yeah, when I lower resolution or bring games back, or like I play with game assets, it's all about trying to reclaim a little bit of lived experience with things that you really have no sovereignty over at the end of the day. Like you can create – like creating is almost the only way you actually have some sort of... this is all you have in terms of your relationship to that algorithm, is kind of the one you have to make real in some sense.

Particular game algorithms contribute to particular experiences, which game companies can regulate access to when they own their underlying algorithms. Game designers and modders like Neilson, in turn, can "reclaim" part of this "lived experience" by implementing these algorithms or archiving them again.

Indeed, for Neilson, game companies can build the lived experience of modding directly into game design.

...that's kind of the new – Roblox might be like the return of that kind of creator economy, where people are memorializing, they're creating all these lived experiences in the game, they're bringing all the facets to life. Just like modding did, but with a more curated set of tools, obviously.

In this sense, the game engine – or the platform of code on top of which games are developed, which determines certain mechanics like physics and graphics – is also a platform for creating experiences:

> *... the [game] engine is a platform to create an experience. And so that like when you have modding, it's like, well, if you have an entire game engine, with movement, and all these things, then you can just easily repurpose a few things to, like, change the entire experience or something like that. I think that's all it is – it's that it's usually just repurposing, actually, rather than building from scratch. You have lived experience in the engine, and you have like a cultural and social community built around this thing. Like already a network to spread this change, right? So I think it's like it's nice, grassroots kind of effect.*

Here we note that relationship between the game engine, the experience that it manifests, and the community that is cultivated around this shared experience. Neilson summarizes: "there's the algorithmic social around the simulated engine that creates the shared reality."

The organizers are also concerned with "lived experience," but it is less the experience of an algorithmic system which could be evaluated to change its design, than a more personal experience that challenges the claims of these systems. For Hamid,

> *... the community has to engage to investigate and understand based on their own lived experiences, that okay, what is it now? ... based on people's lived experiences, you know, you can't fool them.*

Attending to community experience is a way to ground the claims and mechanisms of algorithmic systems in a material way.

For Oscar, the game designers, and the organizers, experience is a way to understand the implications of algorithmic systems. But as Oscar observes, the way that the concept of algorithmic experience can be used is considerably broad. It could be used to evaluate and improve the experience of using algorithms in designing systems or games, or as Neilson discusses, to articulate a community of shared lived experience around specific algorithmic mechanisms. And Hamid's remarks indicate that this community of lived experience can challenge the claims of algorithmic systems.

5.4 Making Visible

And yet, one of the key concerns about algorithms in current theoretical discourse is that they are opaque and difficult to perceive – algorithms affect people without their awareness. Does this preclude an analytic of experience? For Oscar, it doesn't:

> *... is there algorithmic experience when people don't realize it? I think there is when they don't even notice it. They don't have a name for it. But they suffer. They experience it even if they don't know how to call it or they don't even realize that there is an algorithm. ... People don't know how to call it, they don't know there is a system behind, but they suffer from it right?*

This is where the themes of awareness, perception, and visibility come into play. For Oscar the effects of algorithms become sensed and perceptible even when they are not explicitly visible or conspicuous.

For the designers Oscar and Hendrik, the task is to translate this tacit experience into a more explicit "awareness" – a term that is used to describe whether people are aware of how they are interacting with algorithms (i.e., "algorithmic awareness" in Eslami et al., 2015). Hendrik discusses how the algorithmic recommendations of platforms like Amazon are "something that people should be aware of... I think fundamentally, we need to ensure algorithmic awareness." For Hendrik designers can achieve this by

... providing people with good mental models about the systems and what they can and what they can't do. And I think, for instance, a clear terminology is very, very crucial, right?

Here for Hendrik "explainability" (i.e., "explainable AI") plays a role in developing a broader understanding of algorithmic systems. This motivates him to advocate for a more grounded, less spectacular, and thereby accessible alternative to the terms "artificial intelligence" and "machine learning," which would emphasize their critical dependence on data instead of their computational "intelligence."

Provisionally, Heuer proposes the term "data-based automation":

... I'm even entertaining this idea, it's probably completely unfeasible, but to really – I think – the world would be a better place if we would call machine learning "data-based automation". Because it's much clearer what it can do, it doesn't have "learning" in there.

With this definition, Hendrik emphasizes that what is significant about algorithms is not that they learn per se, but the "inferred model" that results:

... algorithm plus data equals the inferred model. And that's actually what we're trying to make understandable. And that's also what we're trying to, to audit there.

In other words, algorithms make inferences and generalizations based on provided data. But as Hendrik discusses with the "explanatory gap," another way to aid in the understanding of algorithms is to explain them in terms of "what is psychologically relevant" to people already, which "relates to how they already perceive the world. And we can use that to make explanations."

Of course, related to whether algorithms are understandable or explainable or people are aware of them is whether they are visible to them in the first place. Hendrik insists that even though the algorithm itself may be black boxed and "invisible," it is this inferred model that should really be the focus of attention. It reveals itself in moments of rupture and breakdown:

> *... it's invisible, and then there's a breakdown, and then it's not. And in a way, I thought, Oh, can we use this? And how do these breakdowns affect the trust in a system? And why can you actually trust the systems when you have these situations where the algorithm goes rogue?*

Here Hendrik expresses an interest in how the breakdown and visibility of an algorithmic system affects user perceptions of it. In particular, he notes that this could be something leveraged by designers.

This interest in "making visible" is something shared by Grayson, Neilson, and Jamie. For Grayson, the function of art in "making visible" finds an analogy in critical Marxist approaches to cinema studies, which acknowledge that "the more invisible something seems, the more actually powerful it is, in terms of what it is trying to do." For him:

> *... really what goes into some of these projects, you know, is like this – trying to upset some of the assumptions about something being apolitical or just like normal or something, you know? To reveal what the status quo is, even when it seems like invisible.*

A regard for human perception is relevant here again, namely in deciding how an artwork should be conceptualized in order to have an impact:

> *... that's like part of the toolkit for sure is like, how it's going to be perceived as like, maybe the only thing that matters, essentially.*

Grayson indicates that a regard for how an algorithmic artwork will be interpreted and perceived is critically responsible for influencing its design.

But the question of how exactly the invisible can be made visible by artists remains open. Is it algorithm logic that should be made visible by artists, or something else? For Neilson there are multiple ways to approach this "making visible":

> *So, I think there's a pressure to make, have artists, like make everything visible, and it was like, trying to resist that of like, 'Okay, am I really just making backpropagation visible to people who understand it?' And like, obviously, when they talk about making visible, it's more like power, hierarchy, and visible things that are kept invisible, but I kind of did resist that pressure to have to make machine learning visible in the ways that we talk about it these days – in the way that there's bias [for example]. And then in that case, my work is a little slippery, where it doesn't try and like pin down a position on where something like a machine learning algorithm has a structure or something like that, but more to play with the kind of the offspring of machine learning, which is, you know, all these new kind of forms of making and doing with them.*

Against the interest in making the structure, bias, or power configurations of algorithms visible as such, Neilson's Crystal Palace project has more to do with making visible emergence and play.

Jamie's concern in "making visible" is as a motivating tool for organizing, on the way to identifying the particular technologies and mechanisms that contribute to social harms:

> *One of the biggest things that I think is so important structurally, in how we make things visible, is when you make things visible, you know where to throw the wrench, you know where to throw the brick, you know what needs to be broken. And I think that was part of what I used as, as like a motivating like tool to be like, we need to learn the word algorithm, we need to learn the word risk assessment, we have to learn the word PredPol, we have to learn the word Operation LASER, because that's what we need to break. And if we can't say it, and we can't name it, we can't break it. You know? And, and that visibility is – that's like, kind of part of visibility.*

Once again, "making visible" has less to do with making algorithm logic visible than developing the language to identify the existence of algorithmic systems and their effects.

For Hendrik, this can be achieved by developing terms like data-based automation. For Grayson and Neilson, it can be achieved through making. And for Jamie, it is a fundamentally community-oriented endeavor – a type of collective study.

5.5 Spectacle and Mysticism

While the notion of a "spectacle" is addressed explicitly by Hamid, it also relates to the use of the word "mysticism" by Neilson. For Hamid, data science and algorithms are used as a "facade" to "mask" violent police practices:

> *... algorithms should also be seen as the process by which they are creating this facade of reform, where they are saying that, 'Okay, well, maybe we can try to take the implicit bias out of this thing. Maybe we can try to kind of just like, you know, just take out that, yeah, well, you know, there's the racist institutions, there's is an inherent bias.' These are the claims that, 'Now we have a computer and computers are race neutral. So all they're doing is that, you know, they're just telling us where to go. And it's all based upon data.' Which is all a false narrative, and absolutely creating an illusion of this, I would say even say, pseudoscience ...*

Thus in the design of algorithmic systems, law enforcement has an opportunity to ask

> *What would be the best way to engage in policing, but to provide and to create this veneer and this facade of science, which then, in a sense, does mask the overt kind of just like, you know, heavy duty policing?*

For Hamid this requires organizers

> *...to look at (algorithmic systems) through that lens of how constantly a spectacle is always created... we can't separate policing, from the larger cultural context.*

While Hamid only mentions in passing that his reference to "spectacle" relates to Guy Debord's *Society of the Spectacle* (Debord, 2012), his own account of a "facade" resonates with Debord's theory in other ways. Both stress how appearances are contrived to create a "false narrative" that distracts from concrete material circumstances, which also incentivizes limited reforms that do not address these material circumstances.

For Neilson, the notion of "mysticism" refers to a similar capacity of algorithmic systems to obscure material processes in a spectacular way. He discusses how he risks doing this in developing Crystal Palace:

> *So like, honestly, most things are invisible in this product at the moment; it really is as as opaque as possible, even slides into mysticism – I'm a little worried right now. So like, I think there is like the both sides of the algorithm argument, which is the making hyper-visible and the making of like mystical, right. And I'm worried I'm actually probably the more on the mystical side at the moment.*

Neilson describes this relationship between the hyper-visible and the mystical as a tradeoff:

> *... the idea is like, yeah, I always want to show them [the audience] kind of how it's sensing. And it's a constant dialogue of like, how much do I – if I show this much, then they're gonna be transfixed on really how it's receiving the world.*

Grayson also touches on this sense of mysticism, indicating that it can be used by artists to their advantage, as an opportunity to then direct attention toward material processes:

> *... it's like being a magician or something, you're like waving it in front of their eyes, like this fun, like, technical, creative thing. And then, and then you get to, like, really talk*

about what you want to do. ... that's like part of the toolkit for sure is like, how it's going to be perceived as like, maybe the only thing that matters, essentially.

Algorithmic mysticism, while it does reveal the underlying material reality of algorithms, can be a way to introduce their implications in a more effective manner. This recalls Hendrik's interest in making algorithms explainable in ways that are "psychologically relevant to people."

In her artistic practice, Danae also views the mystical as more of a tool than an error. She describes how her interest in the mystical stemmed from her dissatisfaction with the types of questions critical AI communities tend to voice:

I've been very frustrated recently, also in the artistic circuits I've been working in ... we will ask research questions that – they seem that people already know the answer beforehand, like, Is AI racist? Yes. Is AI sexist? Yes. 'And that's because the society's racist.' Yeah, that's right. I don't know, research on that. But I like to think of the technological realm as a tool that allows you to eventually ... allows us to access other aspects that are not really accessible through the tools that we have now.

What Danae calls for here is not techno-optimism instead of critique, but an interest in moving beyond tried questions, and understanding what technology can be designed to reveal about experience:

I guess I became an activist because of that frustration – it's because I was tired of, again, of stating the facts again over and over. And that's why I became more interested in the occult ... I became more interested in the nonhuman, and I guess those are questions I would like to explore at the moment, something that is... centered in the human experience, especially in the light of the current crisis, climate emergency.

For Danae, the mystical, the occult, and the nonhuman are ways of centering questions of human experience.

In particular, Danae argues that an insistence on explanation and making visible in artmaking is compromising its ability to ask more interesting questions. With respect to her Twitter bot:

> *It is not me explaining anything. It is the audience who is finding meaning in their own heads... something that plays in the viewer's head, rather than in the work of art itself. But it's not really providing an explanation of things.*

This interest in coincidence and experience could be seen to, as Neilson puts it, "slide into the mystical." But Danae's practice is explicitly focused on materiality:

> *In my experience, there's always, I don't know, a buzzword, which every two years everyone is concerned about. And this in my opinion only contributes to, yeah, to distract from the oldest problem of all, I guess, class struggle, and the prioritization of labor. Things that in my opinion are very material.*

How then, do we reconcile the mystical phenomena of computation with an emphasis on the making visible of its materiality? Danae admits that it presents a challenge:

> *I like materiality, and I like my politics, I guess, based on material issues, but still this paradox, contradiction I experienced by working in the digital... which is also a lot about the unknown. Yeah it's a problem for me.*

For Danae, the mystical is on the side of the unknown, or spontaneous patterns of emergence from hard constraints. She likens this emergent space to a dream space which, despite operating at a remove from everyday behavior and knowledge, has a direct material relationship to the known:

Dreams are a space in which you can access a type of knowledge that you cannot access when you're awake. This is why I guess dreams are so central to psychoanalysis and so important in the most important works of literature as well, when something in literature happens in a dream, it basically represents why this oneric realm that is different from the real one. So yeah, that's what dreams are to me – a space of knowledge, a type of knowledge that you cannot get in the real world.

By playing with the aesthetics of dreams, Danae emphasizes computational aesthetics of emergence, showing how they can take precedence to explanation:

... of course ... I don't want to ... obtain something like the answer to all questions in the world, but it has tended to be a very – I don't know, tender, good experience when I see [the] outputs of the bot. Other people also have reacted to that. And I like it when I think all of that is based on dreams.

This is not to obfuscate or mystify the logical rules that make the dream bot work, but to embrace the mystical as an aspect of their emergent results.

After all, Danae's emphasis on the mystical is always paired with her interest in accessibility. Even if she is not interested in explanation per se, this does not mean that she is not working to make technology understandable and intuitive to others:

... accessibility has always been at the center of my practice, and I guess I am trying to apply it. Now when I teach at art school, I think that by having a material approach to computers, the internet, technology, security, InfoSec, it's a nice spot for learning and knowledge generation ...

Here, by material, Danae means to distinguish pedagogy from more theoretical explanations that do not involve a practice of doing, or in this case, hacking. This ethos echoes Matt Ratto's practice of "critical making" (Ratto, 2011), where making technology is also a means of learning about it, in part because it enables you to understand its emergent properties.

Here again, a kind of magic emerges that Danae attributes mainly to hacker communities:

> *... if you've been to hacker conferences, it's very, very magical, I guess. And it might for some seem a bit outdated, but hacker communities are still very alive. And I think their ethics and aesthetics are something that still interests me.*

For her, the parallels between hacking and magic are generative:

> *... coding is very much like casting a spell, and with the right magic input, you'll be able to obtain a desired outcome ... in the community aspects, and in the anonymity – that was also another parallel I proposed – comparing how [shamans] disguise themselves as an animal to access another state of mind, and how you, in the digital realm also, are able to to shape shift, digitally shape shift. So yeah, I don't think it's everywhere, but it's in lots of places.*

In this sense, the mystical does not exclude considerations of the material. In particular, Danae critiques an account of magic that is associated with identity instead of material inquiry:

> *... that's something that has been bothering me a bit recently, because I think there's kind of trend in general among millennials, Gen Z, is of this interest in witchcraft and magical stuff. And it's strange, because it's not really – I guess because of my own mestiza background of growing up in Chile, where basically magic is basically everywhere, like things you do when you, I don't know, things like, I don't know, making a wish when something happens, or knock on wood when you think of something that you don't want to happen, or love spells, things like that. And it's like very pervasive – it's all over the place ...*

Instead of an identitarian fascination with witch identity, Danae describes the magical according to Sylvia Federici's account of it in "Caliban and the Witch" (Fedirici, 2004), and Frederico Campaña's in "Technic and Magic" (Campaña, 2008), where the magical is something that emerges out of material community relations:

> *... it's participatory, but maybe it's more – it's there in the culture and the community uses for its own benefit.*

Altogether, whereas Hamid, Jamie, and Neilson refer to the spectacular and the mystical as something that obscures the materiality of algorithmic systems, Danae does not view the mystical and the material as necessarily opposed, even while acknowledging the tension between making computation explainable and interpreting it creatively. What, then, is the relationship between the mystical and the material? How is the mystical formed out of the material rules and constraints of algorithms? Is it reducible to them?

5.6 Alternate Reality

One theme raised by the interviewees that addresses some of these questions is that of "virtual reality." For Neilson and Jamie, what is significant about virtual reality is not that it is totally false and contrary to reality, but that it is believable and has a real effect regardless. Neilson namely denounces a hard distinction between the virtual and the real, and tries to produce work that illustrates how this dichotomy breaks down:

> *... when you build these experiences, it's just like trying to untangle or retangle the two – like the virtual-real and the real-virtual, and then having people come into those spaces and kind of realize that impact or something like that.*

Neilson is interested in how virtual environments cultivate real material experiences (recall: "there's the algorithmic social around the simulated engine that creates the shared reality"), and to what extent this undoes a distinction between the real and the virtual. He notes that the precise relationship between the real and the virtual remains undecided: "we don't know how to navigate or parse that space really yet."

Meanwhile, Jamie expresses skepticism about this construction of a virtual reality and its influence on life:

> *... what's really freaking me out a lot lately? Is this kind of alternate reality that's starting to be created ... Sometimes I feel like that's like this other dimension that's being created, kind of like without our permission. So like, all this data is floating around. And the story is being constructed, like Spotify, like constructed a story about you and what your taste was, at the end of the day, at the end of the year, you know, your photos, they'll like, you know, look at all these photos and create a story of a day, you know, and then you start to almost buy it. You're like, "Oh, yeah, that's me. That's me." And instead of – which is the most long, the most difficult thing to do through your life is finding out who you truly are, and what your tastes are and developing your individuality – and now all of a sudden, these machines are creating your reality for you. And a lot of people are buying into it.*

The notion that "machines are creating your reality" implies a type of spectacle that runs against a direct apprehension of materiality; and yet, the epistemic status of this "alternate reality" is tricky because it is derived from data about a person, and is also believable and influential. It is not simply a "facade" in Hamid's sense, but an "other dimension that's being created ... without our permission."

Jamie elaborates this notion by using the term "decontextualization." She describes how this idea emerged as a part of research project called Our Data Bodies, which explores how "marginalized adults experience and make sense of the collection, storage, sharing and analysis of their personal information," as well as the tactics they use to confront these practices (Saba et al., 2016). In the Our Data Bodies workshops, participants are prompted to think about the different devices and methods that can collect data about them, and to reflect on the information that can be gleaned from their appearance and personal belongings, like the contents of their wallets or metro pass cards. Jamie reflects on how the idea of decontextualization emerged through this exercise:

> *What people really kind of started to talk about was this decontextualization of who they were, right? ... So their life is like full of everything, not only their eyes, the way their body works, not only the fact that they bought formula, walking down the street to enter the store, the cameras that caught them – they realize that there's a whole story to even how you got into the store to buy the product: you know, what you look like, how you were even walking that day, what you were wearing that day. And I think what was really great about introducing the different types of technology that could pick up not only like what is in your wallet, your credit cards and stuff like that, but how your body looked, where your positioning was, your phone, being able to geolocate you, is that people started to realize the story behind the event, the story behind who you were, was starting to be decontextualized. And how dangerous that was.*

Decontextualization is a technical practice of collecting partial data about someone, and then subsequently reassembling it to create a new story about them that isn't adequately representative of the context of their life and activity. It reflects the notion of the "surveillant assemblage" (Haggerty & Ericson, 2000): a system for collecting data samples from observed bodies and reassembling them elsewhere as "data doubles."

Jamie describes how an awareness of this process emerged through the Our Data Bodies workshop, stressing the discrepancy between data doubles and personal stories:

> *... afterwards, when we started to talk about algorithms and risk assessments, we realized that this decontextualized data, this data that had no story that was just like one little piece of information of a huge story, was now going to be re-aggregated through an*

algorithm, through a risk assessment, and through a lens of criminalization, or through a lens of, you know, a Target algorithm that was going to start to send you information, and that it was going to tell a story that was no longer yours, that was no longer a story that belonged to you, that had your voice, had your emotion, had your explanation, had your vision behind, your sentiment behind, what was actually happening, and what you perceive was happening. Which is really important when we think about our lives, because, you know, we can interpret people's actions and the things they're doing, and you know, make sense of it on our own. But it isn't until that person speaks about what they were feeling, what was going on in their lives at that time, do we truly understand who they are, and what they were going through, right. And so for me, the Our Data Bodies projects was really about not only getting people to understand the different technologies that were decontextualizing them, or that were breaking them down into data, but then also making them realize that there were algorithms and risk assessments that were re-integrating that or re-aggregating that data and creating a story. And that story was being sold as reality and as truth. And that truth was so far detached from the actual person that experienced something.

What becomes evident through Jamie's account is firstly the way that decontextualization is intimately tied to the politics of self-perception, self-representation, and epistemic justice; that is, who has the authority to mediate someone through a particular lens? Here the issue is not simply that decontextualization misrepresents bodies, but also the way it bypasses their ability to represent themselves. Second, on the other side of decontextualization is a recontextualization that serves the function of "creating a story." It is not simply that data doubles enable decision-making without anybody's awareness; instead, they can be mobilized to craft narratives that manifest as stories to human perception and judgment.

For Jamie, this function of decontextualization in policing is to represent people and communities in a way that justifies particular kinds of treatment by law enforcement:

> *... we get two realities that are being created, right, we get the reality that's being created by either a consumer market that wants to sell you something, and assumes that you're a certain kind of person that needs a certain kind of object, or product, or whatever. But more dangerously, law enforcement is continually trying to legitimize its presence, and its targeting of communities and people, so now they're able to re-aggregate data to create reasoning for them, like you know, reasonable suspicion, probable cause for them to actually target you detain you, shoot you, murder you, push you out of an area.*

What is especially concerning for her is that people have little control over the creation of this alternate reality which is created "without our permission," and also that they can even start to perceive themselves through it, despite what they know about themselves.

For Jamie and Neilson, respectively, decontextualization and the construction of a virtual reality are ways that algorithmic systems are used to cultivate a particular sense of reality. They both address how control over algorithmic processes amounts to control over the construction of this reality, which raises fundamental power asymmetries. Insofar as law enforcement and game developers control the algorithms, they are the ones deciding the types of experiences and realities that people perceive perceive through them.

5.7 Stories

Accordingly, both the organizers and the game designers express an interest in "reclaiming" the lived experiences determined or produced by algorithmic systems. For Neilson, "it's all about trying to reclaim a little bit of lived experience with things that you really have no sovereignty over at the end of the day." For the organizers specifically, but also for Francis, the role of "stories" is key here. As said, Francis expresses an interest in procedural generation in part because of the "storytelling communities" that narrativize procedural generation in creative ways. For him procedurally generated games give "players space to kind of make their own stories." These creative, story-based appropriations of algorithmic processes relate to his interest in the tension between the determinations of rules and their creative appropriation, which he associates with the notion of States designing policies that disadvantaged groups circumvent in turn.

For the organizers, the function of stories is even more explicitly political. For Jamie, a practice of storytelling and "keeping our stories" challenges the ways that law enforcement represents people, communities, and state violence:

> *... keeping our stories is something that we've always fought for historically, in social justice movements, I think primarily because our stories start to expose the various types of oppressions that people are living under.*

And in the context of organizing against the harms of algorithmic law enforcement technologies, storytelling can account for the lived effects of these technologies in practice. As Jamie recounts about the Our Data Bodies project:

> *There was a lot of storytelling that came out of that where especially we were seeing the criminalization of parents, particularly women, and the taking of their children based on different risk assessment models that were just flagging them for ... like a hotline, if someone called in hotline and [said] "drug" or any kind of certain word was used, it immediately flagged that mother to be to be interviewed and to be basically, you know, criminalized and targeted.*

Hamid also describes how telling a story is a critical part of his organizing practice:

> *...based upon community's experiences, that's when we sit down, and you've been a part of those conversations, we're like, 'Now, let's find out exactly what is this the makeup of this new claim? What are the various moving parts? Because you know, if they are claiming this, this razzle and dazzle, so, okay, let's cut through it.' ...So that's when, you*

> *know, we start in a very methodical and practical way to understand [the technology], because then our goal really is, and if we are working to abolish it, then we are not there to investigate it to make it better. We are there [to ask] how do we tell this story to the larger community? And in order to tell this bigger story to the community, the community has to engage to investigate and understand based on their own lived experiences, that okay, what is it now? So in order to find out what is it now, that's when we start using the tools to flip the script where we file for public records – but that by itself becomes a critical piece as well, that, how and what are we looking for? So in that, you know, that unlike a legal process of filing for public records, or discovery, our goal is to tell the story. So it has to be a very comprehensive investigation – it has to – so we start asking for their documents.*

Hamid emphasizes the role of organizing "to tell the story" about algorithmic systems after investigating them in detail and in the context of community lived experiences. Jamie again:

> *...this is where the role of the community comes in. Because in the very first meeting that we had in 2016, where we, you know, we didn't know anything about algorithms, I had no clue what I was looking at. We just broke down a couple articles, we read them in a group. And everybody was like, you know, this is bullshit, they're still gonna come beat us, they're still gonna come beat our asses, they're still gonna come stop us. Like the impact of policing is not going to change. Yes, some people may feel comfortable now about policing, because now police are selling it as like this kind of science that's, you know, clean, and it's cleaned up, and it's more efficient, and it's, you know, it's unbiased. But the community immediately saw through that, and again, we go back to their stories, going back to the stories like who was on the Chronic Offender Bulletin, or who was a Chronic Offender, who was on the list, what areas. If you look at all the the Anchor Points throughout Skid Row historically, those same residential hotels have been*

historically targeted since the 1940s, when there was tenements and they were trying to wipe clean all the black and brown people that were on Bunker Hill, when they successfully did it and pushed them all into Skid Row. Those same hotels were targeted by the Citywide Nuisance Abatement Program. That whole area – they freakin still want to clean them out. They don't want them there, and so they so they police them, so they're hunting them, stalking them, making them uncomfortable, changing the community space so that they can't even survive.

Jamie describes how organizing can center the experience of the community, to "go back to their stories," to challenge how algorithmic systems are marketed as "more efficient," "cleaned up," and "unbiased." Indeed, this is to appeal to lived experience as a way to contest the "spectacle" that is developed to market these algorithmic systems.

5.8 Gamebreaking

The notion of telling stories that challenge the spectacle, alternate realities, or narratives of algorithmic systems relates to a series of themes discussed mainly by the game designers and artists. This theme is more complicated and requires some unpacking – it has to do with the relationship between the determinations of algorithm logic and behaviors that emerge to challenge these determinations. Francis introduces this idea as follows:

... when people – when States and stuff – design policies ... so that only a certain kind of person can benefit from something, or something like these kind of really archaic and convoluted [algorithmic] targeting methods, and, for example, social programs, or something like that – people find ways around it, or ways to evade it and stuff like that. And so this constant frustration of human ingenuity and State control, I think, is also kind of an interesting place where this comes up in the real world.

For Francis these ways that "people find ways around" a system's constraints can emerge out of the constraints of the system itself:

> *I think it just kind of, it's inevitable that these kind of weird edge cases, or unintended behaviors come up. And I guess part of why I like them is because it does happen all the time in the real world. And I think a great deal of effort is, well, these kind of edge cases are, I think, where a lot of kind of human resilience and innovation really comes up, people find these weird ways to make do or change things in a system that is very strongly designed to not allow those kinds of possibilities.*

Francis is thus interested in designing systems that generate these "edge cases" and "possibilities," despite their determinations.

For his part, Neilson distinguishes such "edge cases" that emerge through the constraints of algorithmic systems as either "glitches" or "gamebreaking":

> *... glitches are just more like breakdown in logic of the engine. And usually I think glitches depending are like normally seen as like, especially with like, speed running actually, like benign or like benevolent things, you know, unless it's like a gamebreaking glitch. I think gamebreaking is like a whole different class. And then if it's a gamebreaking glitch, then yeah, I think it is like, 'Okay, well, the developer screwed up, the game's broken, the algorithm's broken. This this is cancerous or needs to be fixed,' basically.*

If a glitch is a breakdown in expected algorithmic logic, whether it is "gamebreaking" depends on how it is used by people. Both are registered according to how people perceive them. Gamebreaking glitches aren't a property of the system itself, but also a subjective expectation of how the system should operate. Neilson gives an example:

> *... the really like liminal, the one that's like right in the middle is like, falling through the map ... like it does break the game entirely. But it's also seen as like this new vantage point where you were never supposed to be there. So that's like right in the middle and then obviously something that just either ruins the sport of competition where something's broken, and you can just cheat, like not good. And there's something that just either entirely impedes progress probably. Like I said, anything that's like benign and like gives you like, a different experience or logic to the game that doesn't normally exist, probably like a plus for people.*

Glitches emerge at the intersection between algorithmic logic and social expectation. Their valuation depends on the type of experience that they manifest.

Certain projects that Francis and Grayson participated in developing together, like Bail Bloc and White Collar Crime Risk Zones, leverage this experience of glitches to communicate arguments. For Grayson:

> *... part of the toolset, or something like that, is also not just the actual image of what you project but also like the act, the implicit act of trespass in what you're doing too, because that's part of what excites people about the project – is not just like the resulting visuals, but the way that we do it, you know, which is essentially like we're making trouble. You know, we're like, transgressing something.*

Like falling out of the map, the negativity of transgression becomes a positive experience, with respect to a new system (e.g., of values, of code). And the fact that a small intervention can operate against the interests of a comparatively big system is something that excites these artists and game developers.

This recalls Neilson's notion of "the barrier between playing and creating," which is enforced by profit-motivated game developers but dismantled by modding:

...i f it's not creatively purposed outside of like the original product, then the [videogame] companies like Rockstar, or like Nintendo – both of them come down really hard, on like, repackaged older versions of games that run on emulators for copyright reasons.

While modders are allowed to manipulate game mechanics as long as they don't create a result that threatens the sales of the original game experience, they nonetheless find creative ways to circumvent this "barrier."

Neilson, Grayson, and Francis – the respondents whose work involves games – are playing with the constraints of algorithmic systems to introduce new spaces and possibilities for creative engagement with these constraints. This type of play has a fundamentally transgressive element – "an implicit act of trespass" (Grayson) – which demonstrates that the experience conferred by a particular algorithmic system (or system of rules) is not absolute. Instead, it can give way to new uses, interpretations, and experiences. It is in this sense that creative play with gamebreaking serves a similar function to storytelling in Jamie and Hamid's organizing.

5.9 Rules

The difference is that gamebreaking and transgressive play, unlike storytelling, take algorithmic *rules* as their object of design. This theme of "rules" is addressed extensively by the respondents whose work involves games: Neilson, Grayson, and Francis. Indeed, for Francis it is precisely "creating a set of rules and boundaries" that gives "players space to ... make their own stories." Creating rules for Francis also plays a critical role in developing games that are persuasive:

... the other thing that drew me to simulations and games is that they have a pretty powerful – a lot of like, rhetorical power, I guess. I thought, the fact that maybe two people could mutually agree on maybe how a system is structured, how it works, but have

different ideas of what the implications of that structure was, meant that if you build a simulation according to those rules, and you showed a different outcome than what they expected, that would be more rhetorically powerful in some way ...

Francis is describing how a set of commonly shared assumptions can be encoded into rules, which serve as a point of departure for a procedural system that reveals the implications of those rules. This has a rhetorical power: "if you both agree on the rules, the outcomes of those rules are harder to refute."

For Francis, when comparing "a simulation on one end of the spectrum, and a game on the other," a game "emphasizes more things like aesthetics, or interactivity, or playability or legibility or like ease of access." Simulations, on the other hand, are more focused on persuasion by departing from a set of agreed-upon rules. For both simulations and games, something is designed to be perceived and interpreted, but the aesthetic qualities and accessibility of the interpretation are emphasized in the case of gamemaking, whereas simulations are concerned with the rhetoric of rules specifically.

Consider Francis' work on the Half Earth game. It is a game that has some elements of simulation, but Francis distinguishes it from his work on creating a simulation for the Institute of Applied Economic Research (Institute de Pesquisa Econômica Aplicada; IPEA) in Brazil:

The thing with this [Half Earth] game and something like the IPEA work, is that here in the Half Earth game, the rules are a lot cruder. A lot of it's just like, I described this, like: if X and Y are true, then this event can happen*, or something like that. And something like the IPEA thing, you don't want to explicitly say like, under these conditions, this event will happen.*

For the Half Earth game, rules encode explicit logics of causality that might not represent reality, but operate instead as a functional caricature of it. For the IPEA simulation, on the other hand, rules aren't designed to encode explicit logics of causality but various relationships already encoded in data or established norms. These relationships may not be legible immediately, as they emerge through the process of the simulation playing out. Francis continues:

> *... so let's say, for example, you're dealing with a simulation of a housing market or something like that. In a game, in a version, like the Half Earth game, you might say, like,* if household debt is above X, or like, household debt is X percent of GDP, then a crash happens. *That's like, you know, I would call it maybe overly hardcoded, the relationship between those two things. And something more like the IPEA, you wouldn't have like a rule like that; it would be more like: well, you have a model of household debt based off of like people going on buying things, or people are buying houses and housing prices, and so on... You're not saying you know,* if x then y*, but based off of these lower level rules, these kind of phenomena emerge without explicitly having to tell it to, if that makes sense?*

At first glance, this distinction would appear to put games on the side of developing determinate rules, and simulations in the business of generating unpredictable emergent phenomena. But this would be to consider only the design of rules; we must also consider the design of the phenomena that emerge from these rules.

For both simulations and games, emergent phenomena are designed through the configuration of rules. And the ways that rules are configured involve different ways of relating to these produced phenomena. As Francis recounts:

> *... like for the IPA, we were kind of interested in designing from the bottom up. We're like, 'Okay, we have, you know, these assumptions or theories about economic like economic behavior at the micro/macro level,' or something like that. And we kind of put those pieces together, and then see if we can reproduce stuff, or what comes of it really, kind of, once we put all those pieces together. Something that's designed for like, a lay audience would probably happen in reverse. So this Half Earth game, for instance. We kind of*

started thinking, like, what kind of events do we want the simulation or the game or the model to produce and then work backwards from there: like what components are important for representing that, like, if we want to show like heavy rainfall and flooding? Well, then we need to incorporate precipitation to it. If we want to show let's see, like the, like regional conflict or something like that, okay, we need to, like, have a module that kind of can allow us to do that, and so on. It's, you know, I suspect maybe Dwarf Fortress was developed in this way.

For the simulation, rules about really existing assumptions, theories, or norms (e.g., "economic behavior at the micro/macro level") are configured "from the bottom up" in an attempt to reproduce other really existing norms as resulting phenomena (e.g., "real data about economic trends"). For the game, there is no need to begin from rules that represent really existing norms, since the main heuristic of evaluation is how the result will appear to the audience. Thus game designers can "work backwards" from the phenomenal results to the codified rules, enabling the design of rules that have little to do with really existing norms.

Francis describes the process of 'working backwards' from phenomenal results to codified rules in designing the Half Earth game:

So it's ... kind of this iterative process of thinking like what kinds of things you want to happen, what kind of variables or conditions will need for those things to like to be able to determine when and if something like that should happen. And gradually, then we come up with like a set of variables like 'Okay, we need to measure like local discontent, maybe public health, if we want to model like the effects of pollution from mining or something like that, or the impacts of heat waves. We need like, you know, emissions because obviously play and the players decisions are gonna affect how many emissions are occurring.' So we kind of end up with a list of these variables, and then some sense of how they relate to each other and how they connect. And then it's, yeah, it's a very *iterative process because as you're thinking of those things like, 'Oh, now that we have these variables, this event could happen.' And then maybe you think of another thing, and that needs another variable. And so eventually, you get to a point where like, 'Okay, there are too many variables, we need to figure out how to cut back.'*

Here Francis gives an idea of how codified rules and phenomenal results are interwoven together during the simulation-game design process. There is a combination of reasoning about really existing chains of causality, how they might be reproduced in a figurative way, and the other possible chains of causality they might entail.

This is where 'working backwards' from phenomenal results to codified rules is also *working back-and-forth* between them and *working transversally* to coordinate their various emergent effects. All to produce resulting phenomena that provide a sense of realism or coherence.

5.10 Rhetoric

Already, we saw Francis identify a kind of "rhetorical power" in the configuration of rules in simulations. He wants "to model as much detail as possible is because that feels more rhetorically powerful," while keeping in mind that too much detail could be inaccessible:

... this is true of probably like any kind of rhetorical medium, is that you kind of have to round off the edges to reach a larger audience, or to make a point more easily digestible, I guess.

A consideration of the rhetoric of simulations and games, and how they should be designed to convey rhetoric, is a key concern here. It evokes Ian Bogost's theory of "procedural rhetoric" in videogames (Bogost, 2007), which Francis notes explicitly.

Indeed, "rhetoric" is is a theme addressed by all of the respondents whose work involves games. Grayson describes a difference between "visual rhetoric" and logical argumentation:

... what visual rhetoric can actually allow us to do is to communicate some ideas, but it doesn't allow us exactly to like, speak about and in the language of the systems themselves. And so yeah, like mounting a simulation, obviously, like kind of allows you to mount arguments about that simulation, or like demonstrate that simulation.

Whereas games are designed to excel at visual rhetoric, the configuration of rules in simulations is to "mount arguments." Francis describes a similar relationship:

So the IPA is more I guess, about developing a theory, right? Like a theory ... you want it to be as simple as possible, right? Like, the more the simple it is, and the more phenomena that simplicity can still explain the more powerful theory it is. And with the Half Earth game, you know, we're not trying to develop a rigorous theory about how the next 100 years on the planet will play out or something like that. We're just trying to help people get an understanding of like, what kind of things we need to happen for the next 100 years to be like, relatively livable. So it's, yeah, it's a very different purpose.

Mounting arguments in a simulation requires a focus on whether the rules are simple, coherent, and elegant, whereas visual rhetoric in games is more concerned with the coherence and elegance of the phenomenal results that players will perceive.

Through the design of game rules, Francis strives to emulate some real world processes, but through the design of results, he is concerned with how these rules will manifest to perception:

... the player needs to feel a sense of agency like that they can actually influence the outcome. Otherwise, it's just a frustrating experience. But the problem is that, you know, a global, like Earth system, climactic systems, and so on, are really complex, and it can feel like you do something, and the outcome at the other end is completely unrelated or completely random. So there are times when, yeah, complexity obscures more than

reveals, and it makes the player feel helpless, like nothing they do matters, outcomes are totally random, even if they're not actually random. Even if there is like a kind of causal chain that you can read out. It's still just might be way too dense and long. So for this particular case, I'm trying to figure out like, well, how much like there's an impulse in me to want to make it like as realistic as possible, like, to try and model as much detail as possible. But obviously, that's not like what we need here. It's not really appropriate for this, for a game of this kind. And so I'm looking for ways to reduce the complexity or actually just like fake it, basically.

Unlike designing a simulation that aims to reproduce real circumstances, the Half Earth simulation-game may sacrifice the complexity and realism of rules for the sake of a coherent result. He likens this to the Borgesian fable:

I guess this is like the – was it that Borges short story – that always gets cited, in cases of modeling and stuff like that On Exactitude in Science*, where, you know, after a certain point, there is such a thing as too much detail, and you kind of lose sight of that, it just becomes too hard to parse. But with the IPA, yeah, we were like, we wanted far more detail than you'd probably want in a game or something like that.*

In game making, Francis can "fake it," or create completely arbitrary rules that produce a result which nonetheless appears realistic. This, of course, comes at the expense of the accuracy of rules, which in turn compromises the credibility of a simulation – 'sliding into the mystical' as Neilson calls it. Francis describes how he has to balance this tension in developing the Half Earth game:

I think the reason why I really want to model as much detail as possible is because that feels more rhetorically powerful. If you fake stuff, and people can just be like, 'Well, that's not how it actually works, you just like kind of said that this was going to be the outcome,' *it's not very convincing, right? So I think the trade off is, is that it does lose a bit of that persuasiveness and hopefully, the parts you kind of leave out but details you leave out, the majority of people kind of do the same hand waving and roughly agree with like, you know, that's a fine kind of thing to leave out. But it's hard to say. Some people will always feel like, 'No, that's the important part that you left out.' But the trade off, I mean, you can get something for like, you know, it's more and more people are going to play it more people will hopefully kind of see the point that you're making. Even if you know you're lacking some of the detail that maybe an academic or someone who has a more, a much deeper understanding of that field might quibble with.*

We already noted how Francis describes the rhetorical power afforded by structural and logical constraints:

The fact that maybe two people could mutually agree on maybe how the system is structured, how it works, but have different ideas of what the implications of that structure was, meant that if you build a simulation according to those rules, and you showed a different outcome than what they expected, that would be more rhetorically powerful in some way ... But ultimately, the constraints of the systems will kind of bind you and push you to act in a certain way. ... And so there's like a kind of logical strength to it, I guess, like, if you both agree on the rules, the outcomes of those rules are harder to refute.

Thus in simulation, there is a fundamental need to let rules simply *play out* from constraints – to have them produce results consistently and without interference. But another type of rhetoric is operationalized through play. For Francis,

... I suspect that the act of playing and interacting, often you're embodying a character, and you're kind of driving the decisions just puts you in a state that's maybe more

receptive to seeing things through a different perspective, or thinking about things differently.

Here embodiment, character, and perspective are all elements of play. They enact a different embodiment and perspective than everyday reality, enabling a kind of defamiliarization that makes an audience more receptive to what they show. Although a game may use logical argumentation or "procedural rhetoric" (Bogost, 2007) to persuade, it involves elements of persuasion that are irreducible to logical argumentation as such.

5.11 Emergent Phenomena

The game designers are concerned with the design of codified rules, emergent phenomena, and the capacity of this relationship to be rhetorical. For them, this relationship *between* codified rules and emergent results is a key site of design. Grayson demonstrates this site between rules and results through his project "why don't the cops fight each other." Grayson's revelation that the police code was not modifiable did not appear by discovering a specific piece of code, but through the emergence of resulting phenomena produced by the game:

There was– I never like– There is no– Well, it's probably possible, strictly speaking, but there's no way to like actually find the code that says, you know, 'don't let this happen.'

Since the code in question is built into the game's logic at such a base level, it is not exposed to those who mod or manipulate the game. For Grayson, this frustrates an attempt to reverse engineering it:

I can't see the thing itself, then we have to go science and like, you know, change one thing at a time until I can like, find the contours of this thing. ... turns out that this is impossible, because the kind of rules that govern the police officers are happening on

> *such a level in like the, you know, the hierarchy of the game engine that it's like close to the kernel, you know.*

But where reverse engineering may fail to reveal the code, the resulting phenomenal effect is no less apparent: the cops simply don't fight each other. The project demonstrates a phenomenon, indeed an impossibility, which emerges through rules that are themselves immutable and inscrutable. For Grayson this ultimately presents as a poetry or rhetoric about the logic of the game and the social circumstances it analogizes:

> *And it kind of makes sense, given the game itself, and the goals of the game and the way the game is set up. To have that be like an immutable situation, that the police, like the game and stuff would not function if the police operate it any other way, you know, and then, of course, that's where, like, some of the poetry or the rhetoric of the project itself, exists as well.*

Unlike a simulation designed to reproduce observable phenomena or causal processes, this rhetoric is not concerned with logical argumentation. Instead, it is concerned with the expression of phenomena produced by logic, the expression of the relationship between codified rules and phenomenal results:

> *With "why don't the cops fight each other?" it's really just through, like, repetition of the thing that I can't do, you know. And like, in that way, it's sort of like, at least, like, draws the outline around the thing that is impossible.*

Grayson's work to "draw the contours" around "the thing that is impossible" involves a logical proof – under all circumstances in the game environment, the cops do not fight each other – but the phenomenal impossibility it demonstrates is irreducible to this logic. Part of the "poetry" of the impossible thing is its inaccessibility, and the unexpected situations it produces in the game engine.

Another one of Grayson's projects, "Usefulness of a Useless Neural Network," also takes aim at expressing a procedural impossibility: the inability of a neural network to solve basic math problems, after it has been trained on a series of examples. Grayson expresses this impossibility through a kind of demonstration that the logic of simulation enables:

> *... with "Usefulness of a Useless Neural Network," I think, just like, through demonstration, you know, through like, this notion of simulating it and running the prediction and like seeing the results. ... it's like less about the drawing the absence and more about like drawing the outline of the absence or something.*

Instead of explaining the logic of the algorithm, Grayson demonstrates what the algorithm is capable of doing – and namely what it is incapable of doing – by letting it play out on its own accord, provided some initial constraints. "Usefulness of a Useless Neural Network" is a simulation in this sense, but one that generates resulting phenomena which deviate from expected norms rather than adhering to them. Once again, the project involves a logical proof (provided examples of math problems, the algorithm does not learn how to solve them), but its poetry lies in the blatant failure of a supposedly advanced learning algorithm to solve a remarkably simple problem.

Grayson contrasts "Usefulness of a Useless Neural Network" to "ImageNet Roulette," a project by Kate Crawford and Trevor Paglen that classifies photographs of faces with obscure and blatantly inaccurate captions. Whereas both projects use algorithms to demonstrate their conspicuous mistakes, the difference for Grayson lies in the focus on data versus code. While "ImageNet Roulette" brings algorithm training data into focus, showing how the determinations of the classification algorithm are limited by the data fed into it, Grayson wants Usefulness of a Useless Neural Network to make a claim about the politics of algorithm code and logic itself.

Rather than supposing that "garbage in" equals "garbage out," the useless neural network demonstrates that the relationship between codified rules and phenomenal results is not so straightforward. Instead, it is manipulated by algorithm logic, which does not become neutral even when its input data does.

Both ImageNet Roulette and Usefulness of a Useless Neural Network involved the intentional design of codified rules and phenomenal results to express something about their relationship. For the former, rules determine results; for the latter, what is at stake is the logic implicated in the relationship between rules and results in the first instance. Grayson compares this focus to historical considerations about the procedures underlying works of art:

> *It's like the drawing on the wall isn't where the art is located. Right. It's more like in the instruction set. And I feel like he [Sol LeWitt] kind of did a lot to complicate this idea.*

But reflecting further on his own work, something other than only the instruction set seems at stake:

> *So with "Usefulness of a Useless Neural Network," like, where's the art located there? I guess I would say it's like, I don't know. It's in the– in the results or something. But then I guess like, those necessarily are visual, they're like, but they're not like, not like aesthetic, they're just sort of data points or something. Or, I don't know we could say that it's like the network itself.*

This point is critical: where is the art located? Not only in the rules – but also in the results, and namely in the relationship between them provided by the neural network itself. This recalls Hendrik's insistence on the term data-based automation: not just the data, not just the algorithm, but the relationship that emerges through the configuration of data and algorithm to produce certain inferences.

Grayson's playful experimentation with algorithmic systems is used to make the "contours" of algorithms visible in lieu of making their actual logic visible. He recounts an experience where some of the more playful results of his artwork were met with disinterest by computer science specialists who were more interested in its technical rules:

> *... they were just like, pointing out technical flaws and what I had done. And I'm like, 'Yeah, but like, isn't that interesting, too?' And, they just didn't really care?*

Indeed, when it comes to an interest in really explaining, understanding, and correcting algorithms, there is resistance to thinking algorithms in terms of their emergent properties, no matter how playful they may be. Attention shifts instead to the logic that produces these properties, as if their true nature is to be found there.

Altogether, artists and game designers may be principally concerned with the appearance and perceptibility of phenomenal results, but their work may also tend to address the role of codified rules in configuring these results. In such cases, it can be a challenge for artists to decide what aspect of the rules-results relationship they should bring to the fore. Should they emphasize the data (ImageNet Roullette), the algorithm, its mistakes (Usefulness of a Useless Neural Network), its aesthetics (Crystal Palace), its coincidences (Danae's Twitter bot), or something else? All the while, we recall how they consider the perspectives of people who will receive their work. This demonstrates that configuring rules and results always occurs with a regard for how they will be perceived by people.

5.12 Summary

Each of the specializations discussed here – designers, organizers, game designers, artists – refers to the idea that something about algorithmic processes could be clarified, if not simply *made visible*, by design. Of all the eleven themes discussed in this analysis, the parameters and politics of *design*, and namely of designing visibility and making visible, are a central aspect.

Each specialization discusses how design is responsible for determining how algorithms are perceived, as well as what algorithms make perceptible, and how their own disciplines can intervene in the formulation of these perceptions. For all of the respondents, the *people* who are subjected to the consequences of these designs, or otherwise strive to understand them, become a key subject of consideration. In particular, the respondents raise the *experience* of people as that which enables them to attend to the consequences of algorithms (through algorithmic experience design), to contest their claims (through stories), or to index the communities that congeal around particular algorithmic logics (through reclaiming the lived experience of gaming algorithms).

Throughout this attention to making the operations and consequences of algorithms visible through design, there is less of an emphasis on the transparency of algorithm logic as such, than on the perceptibility of their phenomenal effects. What is being designed and made visible is not algorithms as such, but something else. To this end, the themes of *rules* and *emergent phenomena* express a relationship between codified algorithm logic and its resulting phenomenal effects, which is a relationship that becomes an object of design and critical inquiry. The work and expertise of the participants is grounded in an attention to the relationship between algorithmic rules and phenomenal effects, which becomes a medium for design, critique, and artmaking in its own right.

This relationship between algorithmic rules and phenomenal results can be designed, mobilized, or confronted in a number of ways. One way is to use it in the service of *rhetoric*, where the artists and game designers experiment with the persuasive possibilities and limits of phenomena that are articulated out of algorithmic rules. Here artists and game designers participate directly in the configuration of the relationship between rules and results, examining and altering its reception by their audiences. Another way to confront the relationship between algorithmic rules and phenomenal results is through *stories*, which are not directed at redesigning this relationship so much as contesting its rhetoric and claims. The organizers focus on telling stories against what algorithmic systems show (although they may also mobilize these algorithmic appearances in the act of storytelling – demonstrating correlations between what algorithmic systems show and what the police do, for instance).

Lastly, *gamebreaking* is another kind of intervention into the relationship between algorithmic rules and phenomenal results, which concerns an "act of trespass" (Grayson) with respect to existing algorithmic logics. Gamebreaking in particular shows that the significance of algorithmic rules is contingent on subjective perceptions and patterns of expectation. What qualifies as gamebreaking or transgressive is not necessarily quantifiable with respect to a system (i.e., in terms of a disequilibrium), but encourages us instead to attend to subjective expectations and beliefs about what a system should do. Artists can, in turn, take advantage of these expectations and beliefs to articulate rules in a way that surprises, provokes, or draws attention. This raises the theme of *spectacle* as those algorithmic phenomena that risk concealing the material process of their production, deceiving subjects with a false or partial representation of reality.

Altogether, insofar as the designers, organizers, and artists attend to deception, they conceptualize it as something more than concealing code under an obfuscatory facade. While Hamid is concerned about the *spectacle* that algorithmic systems present to distract from longstanding material practices, he and others also acknowledge that this spectacle is irreducible to an act of concealment. In particular, the power of these systems lies precisely in what they reveal about existing circumstances – to crime analysis, to the public, or to ourselves – by "decontextualizing" (Jamie) these circumstances and arranging them in a new way, like an *alternate reality*. Thus conceived, the notion of algorithmic bias does little to account for how data is recontextualized to suit a policing agenda, construct a crime narrative, or cultivate perceptions of crime and space. It becomes necessary to look beyond the logic that decides particular associations of data in Palantir Gotham, for instance, and toward the higher-level phenomena that algorithmic systems are designed to articulate to certain ends.

For their part, some of the artists conceptualize the spectacular appearances of algorithms as a kind of mystification. While Neilson fears that a failure to center the material reality of algorithm operations could give way to mystifying them, Danae is interested in the mystical as a generative capacity of algorithms. But both acknowledge a tension between explaining algorithms on the one hand (the material), and using them to dynamically generate phenomena that are open to interpretation on the other (the mystical). They are both sensitive to the types of questions that are dominant in algorithm criticism, which involve a resounding demand for clarification and revelation. Danae responds to this demand by positioning mysticism as less a distortion of the truth than a power of conjuration, which can be used to artistic ends, or perhaps to more nefarious ones as well. Neilson wants to avoid the trite solution of simply making an algorithm's implementation details visible, and his work focuses instead on experimenting with its generative effects. Indeed the 'slippery' nature of embracing mystification in favor of revealing materiality may not be harmful. Instead, might it be a tendency to engage with paramediation as an object of design?

Hendrik's appeal to designing algorithmic systems from the ground up to enhance their accessibility evidences an attention to algorithm appearances as a fundamental aspect of their operations. For its part, Oscar's algorithmic experience design points more broadly to the need to understand how algorithm appearances are designed to evaluate their social effects. And Hamid and Jamie, while regarding the design of these appearances as more of a threat than an opportunity, would be the first to insist that algorithms are designed explicitly with their appearances in mind. Lastly, through participating firsthand in the implementation of algorithms, deciding how they arrange appearances, and questioning how they are perceived in relation to other social issues, the artists and game designers are uniquely positioned to apprehend paramediation as a design medium. In the next section, I discuss how I implemented a game in order to develop this apprehension firsthand.

6 Egrecorp

Game designers implement procedural rules that arrange appearances in coordination with player actions, in such a way that they can emulate phenomena and elicit particular interpretations. This suggests that, by making a game, I could also account for the process of designing paramedia firsthand. But is it also possible to demonstrate the theoretical, political, or practical implications of paramedia to an audience that plays with them? This was the motivation behind developing Egrecorp, a game designed to simulate certain types of paramediation and to narrativize their theoretical significance. Making the game demonstrated the specific challenges involved in doing so, as well as particular aspects of paramediation that helped to refine the theory. Egrecorp was exhibited at the exhibition "Assemblies, Swarms, and Intricate Webs - no solidarity exists in a social void,"[8] and it remains playable online at egrecorp.com.

6.1 Motivation: Games and Governance

The initial design and concept for Egrecorp was motivated by the idea that videogames involve design conventions that are also employed throughout instances of algorithmic governance. Through a "research-creation" practice, I could experiment with and illustrate the functional and aesthetic resemblances between these systems (Chapman & Sawchuk, 2012). To be sure, from the perspective of gamification, the idea that games and governance systems share design conventions is not new. But while gamification refers to algorithmic systems that are overtly aestheticized as games, the algorithmic systems I analyzed as case studies are designed in ways that exhibit game design conventions in more innocuous ways. While not considered an instance of gamification, for example, TikTok is designed to reward users for certain activities, like posting content, with notifications. As well, Palantir Gotham could be conceptualized as a sandbox for playing with combinations of crime narratives. The point here is not that these

[8] https://www.akademie-solitude.de/en/event/fragile-solidarity-fragile-connections/

systems are literally games, but that they involve design conventions normally associated with games exclusively.

This suggested that game design might be able to demonstrate innocuous design conventions involved in algorithmic governance more clearly. By using game design conventions as analogies for those involved in algorithmic governance, game design might be able to make mechanisms of algorithmic governance more obvious or accessible. As well, by identifying design conventions that are shared by videogames and instances of algorithmic governance alike, I could demonstrate how paradigms of computation normally associated with videogames are also, in fact, involved in algorithmic governance as well. This is significant because, while algorithms are often conceptualized as texts (as recipes, as rhetoric, as code of law), as images (as representations), and cinematically, algorithmic governance in critical scholarship is rarely ever acknowledged to involve the dynamic and interactive capabilities of other computational media like videogames.[9] Videogames offer a clear example of these capabilities, which go beyond those of texts, images, and films. Indeed, if algorithms govern more like videogames than other media, then how should that inform critique?

To be clear, in popular imaginaries, videogames themselves are often associated with the notion of codes of law that govern player activity. This neglects all the ways that videogames are designed to configure perceptions, and to orient behavior without regulating it like a legal system or a 'choose your own adventure' text. Game mechanics in Egrecorp were inspired by horror games for this reason, as the horror genre more clearly involves the design of procedural rules to elicit particular sensations, usually to instill fear or surprise, rather than controlling player actions outright. For horror games, data about a player's behavior is processed by simple algorithms to deliver audiovisual stimuli that motivate particular actions. Algorithmic phenomena emerge

[9] On the other hand, there are many instances of game studies that conceptualize games as approaches to criticism and allegories for governance systems (for notable examples see Galloway, 2006; Wark, 2006).

through paramediation: arrangements of phenomena that appear to stalk players or haunt their behaviors, without any distinct visual representation.

For two key reasons, the design of Egrecorp was also motivated by the conventions of decentralized autonomous organizations (DAOs). DAOs are systems of code-based laws that regulate user activity, which users can vote to change, typically to earn points that increase their voting power. The first reason for invoking the mechanics and language of DAOs is that they epitomize the idea that algorithmic governance can be managed like a code of law. Interest in DAOs rests on the idea governance can be refined and optimized by implementing transparent code that legislates activity: permissible actions are determined absolutely by a system of rules that is visible to everyone. The second reason is that despite this insistence on regulatory codes, DAOs are everywhere designed like games. DAO designers are constantly concerned with how to prevent 'gaming the system' of votes, and how to incentivize desirable user behavior through 'incentive design.' Together these aspects of DAOs demonstrate how game design conventions are involved in governance systems. They also provide an opportunity to explicitly engage with the notion that code is law.

6.2 Design Process

To begin designing Egrecorp, I set out to implement a simple algorithmic system for each of the functions of paramediation that I identified in Chapter 3. To do so, I used "game sketching" to outline, draw, and prototype game mechanics and narrative elements that would suit this task (Westecott, 2020). I decided to sketch *toy algorithms* that could be made into individual levels or game environments that players could interact with. I also sketched overarching game mechanics that would be present throughout all of the toy algorithms, and which would emphasize the relationship between *codified rules* and *phenomenal results* that I identified in Chapter 5. Namely, I aimed to demonstrate that, irrespective of the rules that are codified in algorithm logic, phenomena that emerge from this logic are designed in a way that is irreducible to it.

The game mechanic that I sketched in an attempt to achieve this is called a *Debugger*. In the current version of the game, it is an item that players equip at the beginning of the game, which allows them to 'debug' various objects in the game environment – to see and edit their underlying code. I sketched a number of ways that debugging could operate: by visualizing flow diagrams, or by displaying code that could be edited. As I began to implement the Debugger mechanics in practice, I designed a system of *code blocks* that players could drag and drop to change the code of the objects that they are debugging. I took this approach because it avoided requiring players to actually have to program, and because it was more tractable to design determinate procedures.

The initial plan to implement a toy algorithm for each function of paramediation shifted as I tried to simplify this task by focusing on the four types of paramedia represented in the case studies: roughly, images, causal events, narratives, and synthetic worlds. Accordingly, I sketched a game level or world for each of these types of paramedia (described later below). Players would enter into each of these worlds and witness corresponding considerations of paramediation. I sketched these worlds as dystopian environments where inhabitants were grappling with the consequences of paramediation. The game narrative was that each world had implemented a type of governance that was failing to operate as intended. This was because the governance systems used paramediation, which its subjects failed to perceive from their vantage that "code is law."

Game development took roughly four months. The beginning stages of development involved experimenting with programming libraries that would be conducive to the sketches I was developing. I landed on developing an entirely web-based game because I considered how it could emulate the design conventions and aesthetics of video games and basic interfaces at once. I also thought that a web-based game would be more accessible than one that had to be downloaded or installed. I used the JavaScript software library React Three Fiber, which combines a library for 3D called three.js and a library for webpage state management called React, and developed the game on a Node.js stack. At a certain point in development, I also used Auth0 to handle player save states. This, however, required players to create an account and log in, and was not conducive to making the game accessible with a 'pick up and play' design. Developing a game engine with React Three Fiber and Auth0 that could be used to develop additional features took at least a month.

The remaining phases of development involved creating a number of levels and interactive experiences based on my sketches of toy algorithms, which were all experiments in explaining certain aspects of paramediation. After developing each of these experiments, I played through them myself in an attempt to get a sense of what they showed and how they felt. Some experiments would seem functional as soon as I finished developing them, but when I came back to them later, I noticed that they were challenging and difficult to use. Along the way, I also invented new strategies for simplifying the game mechanics, which I had not incorporated into earlier experiments. This motivated me to update these earlier experiments accordingly.

Throughout development, I drastically changed the amount of text that the game depends on to guide players and explain concepts. Whereas too much text could risk distracting from the game mechanics, too little explanation could leave players in the dark about what was going on in the game. It proved to be a challenge to tell the game narrative and explain the meaning of game mechanics. For this reason, I decided to tell the story by way of *code blocks* that were attached to various non-player characters in the environment. Players could debug these non-player characters to view their *code blocks*, which provided information about the story and the game environment.

Figure 6.1: *Code blocks* in *Egrecorp*. The *Debugger* is also visible in the bottom right.

6.3 Game Playthrough and Feedback

The current version of Egrecorp is a website that begins with a landing page that says "CODE IS LAW – WELCOME TO PARADISE." This page is supposed to indicate to users that they are entering an environment which is supposedly utopian, where code is law. This basic cue appeared to be effective according to feedback from playtesting. Oscar describes his experience after playing the game:

> *I had that feeling of tension, and discovery, of even apocalyptic games... that use this metaphor of making the player feel inside this apocalyptic world, and there are new rules going on, and you have to play around with these new concepts and the rules of this new world. And I think that's a positive concept of the game that made me feel in that sense, in that world, and made me raise the questions and awareness of also what we are doing in terms of in a digital society, and also digital democracies, and also digital governance.*

Clicking the screen plays an introductory sequence that indicates a concern with "algorithmic fiction." I chose this term because fiction connotes a type of human design that does not reflect reality absolutely, but which is nonetheless believable. I tried to use the term as a way to introduce paramediation: this is a digital society concerned with the fictions that algorithms are producing in some way. When asked about the function of this term in the context of the game, Grayson reflected that it suggested how the game illustrates "the space between what these things are purported to be and what they actually are." This is certainly the main implication of the word "fiction." At the same time, I used the term to imply more of a concern with uncovering the material reality of algorithmic systems, rather than with the effects of these fictions in their own right.

In the next step of the game, players are told that they should download a "Debugger" that they can use to find the truth behind algorithmic fictions. They can then use the Debugger to inspect various objects to reveal the code blocks involved in them. This mechanism was supposed to show to players that there is a gap between algorithmic rules and their phenomenal effects: even when the rules are absolutely visible, their effects are irreducible to them. In an early playtest with Neilson, where I designed the *code block* mechanism to manipulate the parameters of voting simulations, Neilson reflected that he was using a "trial-and-error" approach to configure *code blocks*, rather than inquiring deeply into their meaning. This motivated me to implement less sophisticated algorithms (the voting simulations involved many parameters and variables that were esoteric to players), and to design puzzles for *code blocks* that were less open-ended – putting the user "on rails" as Neilson recommended.

The next part of the game is a room containing a black box in a 3D environment. My initial goal was to use a literal black box to show that, despite what users find inside of algorithmic black boxes, the arrangements of phenomena that result from these algorithms are irreducible to the contents of the black box; that is, irreducible to code. This was incredibly difficult to demonstrate in a game environment, since any incoherence between codified rules and their resulting phenomena was more likely to be perceived by players as a glitch, or as their own mistake, rather than expressing a discrepancy between algorithmic code and phenomena.

Nonetheless, players did appreciate the use of the black box metaphor. For Oscar,

> *... definitely there is a very interesting concept there. Going into the black box, exploring the black box, everything being black, everything being hard to discover, which is definitely a very powerful metaphor.*

The issue here is that game failed to intervene into the existing concept the algorithmic black box. Rather than portraying the black box metaphor in a new way, the game simply seemed to make use of its implications.

There are five worlds that the user can access by loading them into the black box. Each world is designed to demonstrate a different aspect of paramediation. Each world also illustrates a different voting mechanisms of DAOs. Players often asked me if the DAO mechanisms included in the game referred to actual DAO mechanisms. Some players thought that the game made an explanation of these mechanisms more accessible, whereas others felt that, without an adequate introduction to the terminology, the inclusion of these concepts was potentially more confusing.

This reflects a key tension of Egrecorp's design. On the one hand, the game is designed to make certain concepts and issues of algorithmic governance more accessible. But on the other hand, the game is explicitly trying to demonstrate that algorithmic systems which advertise accessibility and inclusivity may in fact be very inaccessible and exclusive. And it does this in a satirical, playful way, which succeeded in causing many players to reflect on these issues and laugh about their irony. Egrecorp uses the concepts, mechanisms, and terminology of DAOs to do this: throughout the game, players face excessive gatekeeping mechanisms, democratic voting systems that create unusual hierarchies, and esoteric language that describes why these mechanisms should exist. And this live demonstration and satire of DAO complexity and inaccessibility was, unsurprisingly, complex and inaccessible itself. Oscar summarized the advantages and disadvantages of this:

> *... Sometimes I felt lost, in terms of the metaphors that the game was using. Nevertheless again that's part of the sense, again, of making you feel, 'Wow, what happens with someone that doesn't even know what's going on, that doesn't even know what a computer sometimes is, or maybe they know what it is but they don't know how to use it, but they don't really understand what is debugging, what is loading, what is opening, what is bots,' or something like that. It made me raise awareness of what is the society that we are building, maybe leaving a lot of people behind as well. Or not allowing them to participate in this society anymore maybe.*

Egrecorp's satirical element and complexity would be a success if the purpose of the game was only to illustrate this inaccessibility. But it is a limitation insofar as the purpose of the game is *also* to make issues in algorithmic governance more accessible. It should not have been designed in an attempt to do both at once.

6.3.1 Combat-Based Voting

The first of five worlds that players can load into the black box contains multiple rooms of distorted textures generated by an I2I algorithm, and floating faces that players can interact with. Its main feature is a combat-based voting system, where players can vote to change the DAO's laws by beating opponents in a first-person shooter game.

Players found this world, and particularly the combat-based voting system, to be the most effective at demonstrating one fundamental aspect of paramediation: that algorithmic governance manifests in other ways than regulatory codes of law. The combat-based voting system was, for nearly all players, the most memorable part of the game. This was not because the combat felt exciting or interesting in itself (in fact it was unclear to many players that combat was going on), but because of a text-based interface that explicitly told players that this was what was happening. The text tells the player that combat-based governance is using a "lazy" voting system, a DAO design principle where any user can propose a law/amendment, and the law will pass if nobody vetoes it.

Another part of the first world is a "bot generator" room where players can compose code blocks together to create a "bot." *Code blocks* are either "normal" or "abnormal," and the game indicates to players when they are using too many of one or the other. Thus players are instructed to balance between normal and abnormal code blocks to create a bot that is sufficiently normal to be believable, but sufficiently abnormal to go beyond duplicating existing data. Grayson remarked that visually rendering the process of making a bot, along with various aspects of algorithmic governance, made "the machine look so much more absurd."

6.3.2 Value-Based Governance

The second world in the Egrecorp black box is an undersea network of pipes containing artworks. When players try to look at the artworks, they are told that they must download a "Valuator" in order to do so. When equipped, the Valuator shows players the value of artworks and other users in the environment. Players can then click on these objects to increase their value. If the object has a high enough value, players will subsequently receive "Self-Value" as a reward; otherwise they will be penalized and Self-Value taken away. Players can also add or subtract value from pipe valves, which are labeled with various governance proposals. These mechanisms together are supposed to illustrate a "holographic" voting system, where DAO users can stake points on proposals that they expect to pass.

Toward the end of this section, players move through a labyrinth which they are warned has high levels of "Anti-Value." The Anti-Value increases and decreases depending on where they move, and it increases if they stand still. This mechanism did not seem prominent enough for players to understand specifically what was happening. Instead, this part of the game was overwhelming, which could be a feeling evoked by algorithmic systems more broadly. Oscar said he

> *... felt sometimes in the final part of the labyrinth a little bit exasperated on not knowing what to do, and somehow that feeling is ... what these kind of systems also create in our lives, in our everyday decisions ...*

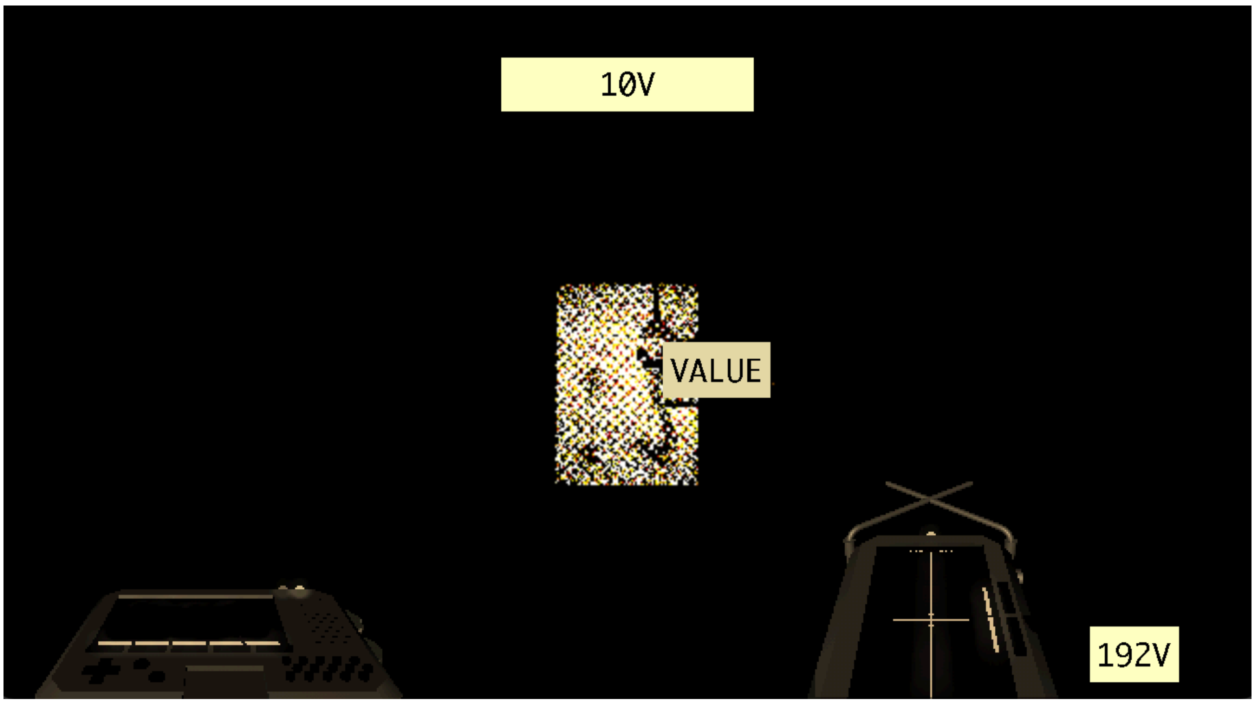

Figure 6.2: The *Valuator* in *Egrecorp*. The player uses the *Valuator* to check the value of an artwork

6.3.3 Filtering Algorithmic Percepts

The third world is a castle, where players must download a lantern called an "Enlightener" in order to view the environment around them. Players can add "data sources" to the Enlightener that enable it to reveal, or otherwise filter out, additional things in the environment. The purpose of this mechanism was to demonstrate in an explicit way how the graphical environment is constructed out of data and algorithms. Some players found this to be a memorable metaphor for demonstrating how algorithmic systems operate.

The castle also contains a "Magic Mirror" and a "Scripture." The Magic Mirror reflects a player's inventory back to themselves, and labels them with a "Mark." The Mark does not do anything except for indicating to players their social status and worth in the castle. The Scripture contains *code blocks* that determine the laws of the castle. The laws do not only enable or disable certain actions, but also alter how the castle is perceived: they change its appearance, its behavior, and its story. Both of these mechanisms were supposed to illustrate how code-based rules can be used to alter a perceptual environment, rather than only constraining activity directly.

6.3.4 Automated Democracy

To enter the fourth world, players have to complete a CAPTCHA test. The CAPTCHA evaluates whether players can "discriminate" between normal versus abnormal *code blocks* and users. When complete, players are told to give the CAPTCHA permission to track their information. Players are then admitted into a world, in which they are assigned to a "Family." A Family is a regional level of governance, which can vote collectively to influence a higher level of governance. Players can be assigned to another Family depending on their actions, for example by going to work (on an assembly line for removing bias from black boxes).

In a room called "Governance Engine," players can view all of the code blocks that determine the laws of this world. One code block is "Autodemocracy," which disables players from voting manually, and instead estimates their vote based on their behavior and Family. Players can thus vote indirectly by adopting certain behaviors. Effectively, governance becomes an incentive rather than a coercive system. This mechanism was not clear to players, and perhaps more guiding text would have helped here.

6.4 Summary

The current version of Egrecorp succeeds in using DAOs as a narrative frame to provoke inquiry, but it fails to draw attention to the algorithms involved in arranging phenomena. While it implements the aesthetics of horror games and some simple mechanisms, it does not use any sophisticated, *noticeable* algorithms to respond to data collected about user behavior. What does this mean, and why is it significant? Egrecorp uses simple algorithms to respond to user interactions with arrangements of phenomena: clicking something or moving an object elicits immediate feedback, or engaging in a certain virtual activity changes the player's score or classification. Because of the simplicity and linearity of these algorithms, nothing seems particularly "algorithmic" about them – they are like the simple algorithms that Oscar describes as not being "experience worthy," such as pressing a key on a keyboard. While paramediation is involved, it is not a particularly complex kind, and is not likely to provoke direct attention.

Second, while earlier versions of Egrecorp involved more user interface aesthetics, blurring the lines between conventions of games and other algorithmic systems, the current version sacrifices these aesthetics in favor of a more typical 3D game environment. The reason for this is that it was more tractable to implement a standardized 3D game environment rather than implementing an environment of user interfaces, whether manually designed or randomly generated. Currently, Egrecorp uses abrupt, rather than gradual changes between 3D environments and 2D UI, which does not blur the lines between the conventions of games and other algorithmic systems, and instead seems to reinforce them.

Another limitation of Egrecorp's current design is that an overall emphasis on the game's aesthetics and narrative detracts from an illustration of algorithmic behaviors. Altogether, players focused their attention on the aesthetics and content of the game environment. This suggests that the process of paramediation, or the algorithmic arrangement of appearances, can recede from apprehension the more that its effects are coherent and engaging. Accordingly, I found myself trying to interrupt this coherence while developing Egrecorp, trying to play with game design conventions and algorithmic mechanisms to interrupt their familiarity and coherence. But this turned out to pose a unique challenge: interrupting conventions of paramediation in games can render them unplayable and unusable (In contrast, for films, incoherence and interruption does not prevent the film from continuing to play itself. Interpretation may break down, but no user response is required for the phenomena to progress in any way. Games, on the other hand, demand continual action.)

This raises one of the greatest limitations of Egrecorp overall: it was demanding, both in action and interpretation. Players were made to navigate through a relatively complex world of interesting graphics and interactive conventions. This required a great deal of mental investment and attention, which proved to detract from observations about more banal algorithmic mechanisms. A more effective version of this project might leverage more banal aesthetics that more clearly involve sophisticated algorithmic responses to user behavior.

Egrecorp is also reliant on text to explain to players the context and narrative of the game, the game mechanics, and their significance. As said, throughout the game development process, I experimented with different amounts of text and different approaches to introducing the game with text. In early playtesting, players would be initially shown a webpage with descriptions of the fictional purpose behind Egrecorp, and made to read fictional arguments between its members. This was very demanding, and did not help players to understand the game mechanics. The current version of the game uses text throughout, which for Grayson was a positive element, while for Oscar more text would have been useful to provide players with guidance. Altogether, while one ambition for Egrecorp was to use game mechanics to illustrate aspects of paramediation, text was still indispensable for contextualizing game mechanics among theoretical issues and social concerns.

Egrecorp was ultimately effective at provoking inquiry about DAOs and digital governance systems, but this tended to require existing knowledge from players about the implications of these systems. Players with background knowledge of this field responded most to the combat-based voting system. Here, a distinction between games and governance systems was blurred to generative ends. Players took time to read the descriptions of the voting regimes in each world, and sometimes asked if the names of the voting systems were real and referred to real practices, which they do. The game did not seem explicitly able to demonstrate the limitations of each voting regime or how it relates to a theory of paramediation, but it did provoke some general engagement with these ideas. This suggests that providing more explanations and descriptions could help to engage with these ideas in more depth.

Overall, players expressed that using a game to make considerations of algorithmic governance more accessible was a compelling idea that warranted further consideration, and that the game was exciting and entertaining in this regard. Without being told that the game was specifically about algorithmic governance as such, Oscar reflected that it could be helpful for demonstrating related concepts:

> *The idea of a game in which people can explore stuff, I think that is definitely, I think that is very valuable to teach about certain concepts, and to make people feel somehow where we are in terms of algorithmic governance...*

In this sense, Egrecorp was a proof of concept that was limited by involving too many game mechanics, narrative elements, and interesting aesthetics, rather than a process that explicitly guided the player through what was going on. This demonstrated the potential of the medium for the task at hand, but did not successfully isolate a specific theoretical issue or social concern for critical reflection.

7 Discussion

7.1 Analysis

A theory of paramediation addresses the capacity of algorithms to arrange appearances according to data, and it encourages us to attend to how these arrangements are designed. While the case studies addressed some applications of this capacity from the vantage of existing algorithmic systems and their basic design considerations, the interviews examined how specialists describe the ways that they and others perceive these systems, and the game design process investigated the process of developing such a system firsthand. In this final analysis, I address some of the considerations raised by these three phases of the study, before addressing some limitations of the study overall, as well as its further theoretical implications.

7.1.1 Approaches to Paramediation

The case studies demonstrate the ways that algorithms configure how various kinds of media appear, from images (I2I), to events that evoke a sense of causality (TikTok), to narratives (Palantir Gotham), to fictional worlds (Dwarf Fortress). Together, the cases demonstrate some common ways that paramediation becomes *an object of design*. This generally involves 1) a qualitative evaluation of the phenomena produced by algorithms, or 2) a quantitative heuristic for measuring whether these phenomena realize a desired effect. Moreover, qualitative judgments about algorithmically generated phenomena may be formalized as quantitative heuristics – depending on the task at hand and the advancement of approaches for responding to it. In either case, the object of design is what algorithms make perceptible. This demonstrates paramediation as an area of research that is being developed through the formulation of concrete design principles and strategies.

In image-to-image translation (I2I), for example, synthetic images can be evaluated by algorithm developers in a qualitative and informal way (see for example Yin et al., 2017), and then subsequently by crowdsourcing quantitative data about how these synthetic images are perceived qualitatively by a large group of people (Zhang et al., 2018), and then again by formulating quantitative heuristics according to this data (ibid.). This demonstrates that, even in cases where paramediation is automated, there is some dependency on qualitative judgments and evaluations. And yet, the generalization of methods, techniques, practices, and principles in I2I research suggests that a computational science of particular paramediation functions is being developed, where qualitative observations about the effects of particular algorithmic implementation details become formalized as 'rules of thumb' or even mathematical principles for configuring certain appearances (see for example Lala et al., 2018). Research continues to refine how particular features and settings of hyperparameters translate to particular phenomena, aesthetics, and subjective perceptions.

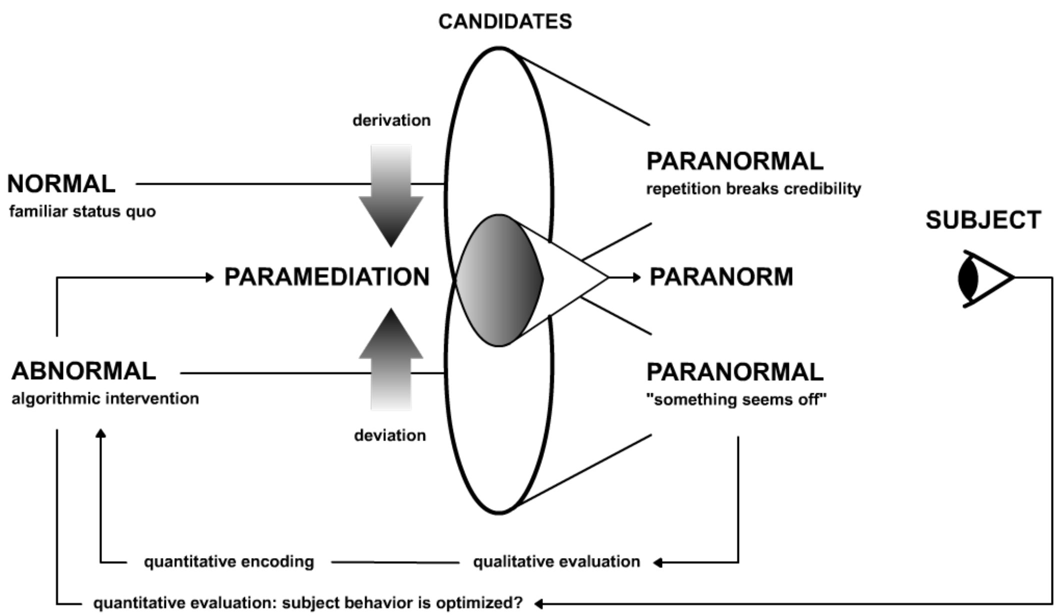

Figure 7.1: Process of paramediation: Given normal and abnormal features are reconciled by paramediation to generate candidates. Candidates are filtered according to the effects that they have on a human subject, which can be evaluated quantitatively by measuring the subject's actions. On the other hand, if candidates convey paranormal sensations, a qualitative evaluation may be necessary to identify this and update paramediaton accordingly.

The algorithms used in TikTok, for their part, can arrange the appearances of user content according quantitative heuristics, which optimize for certain behaviors like user engagement (to be sure, qualitative observations have a critical role to play in this configuration as well, insofar as user perceptions of content inform their interaction with the platform, providing TikTok with data that informs what content to recommend). In this type of approach to paramediation, there is a content-agnostic optimization of appearances, which can result in unintended side-effects – like the optimization of shocking, disturbing, and discordant aesthetics that nonetheless increase engagement (Bridle, 2018). Here, the use of quantitative heuristics to shape qualitative content results in manifestations of the *algorithmic paranormal*, where algorithmically generated paraesthetics exhibit a sense of coherence and structure (determined by guiding parameters or heuristics) that is not informed by direct human design. Notably, unlike the case of I2I where the goal is to synthesize particular images according to certain criteria, perceptions of paraesthetics on TikTok does not necessarily impede its task to optimize engagement.

Meanwhile, the design of the algorithms for articulating crime networks in Palantir Gotham is based primarily on qualitative judgments – about the capacity of particular articulations of data and information visualizations to elicit a sense of coherence in the production of crime narratives. The same is true for Dwarf Fortress: quantitative heuristics do not play a significant role in optimizing algorithmic appearances, and it is up to the game's developers and players to evaluate them.

7.1.2 Normal and Paranormal Media

One design paradigm present across all the cases is that of balancing between, on the one hand, the given data or algorithm logic that informs algorithmic operations, and on the other, the resulting articulation of this data, which derives and deviates from it. I2I illustrates this most clearly, as synthetic images must reflect the source images while also deviating from them sufficiently (although not excessively). TikTok demonstrates this as well in the balancing between the consistency and novelty of recommended content. Palantir Gotham, for its part, balances between the coherence of crime narratives and the spontaneity of investigative discovery, which avoids an impression of determinism. And Dwarf Fortress balances between procedural coherence and generative spontaneity as well. We can describe this balance as one between the normal and the abnormal, to arrive at a paranormal distribution:

	I2I	TikTok	Gotham	Dwarf Fortress
too normal	the synthesized image simply reproduces the source image	content is too repetitive and boring	crime narratives are limited entirely to the display of crime data	the game plays out deterministically without spontaneity
too abnormal	the synthesized image appears too different from the source image, or algorithmic influence is apparent	content is too random and not suited to user interests, or algorithmic influence is apparent and creepy	crime narratives have no coherence or involve apparently arbitrary principles of determination	the game plays out arbitrarily without determination
paranorm	the synthesized image is sufficiently similar to and different from the source image	content is sufficiently consistent and spontaneous	crime narratives are coherent but not determined absolutely by algorithmic rules	the game plays out in a consistent and spontaneous way

Table 7.1: Normal and paranormal phenomena exhibited by each of the cases

Another way of characterizing the balance between the normal and the abnormal in terms of the extremes "overfitting" and "underfitting": circumstances where machine learning algorithms derive parameters that correlate to their input data too totally or not sufficiently. Machine learning must always balance between normal determinism and paranormal stochastic processes, with respect to input data. However, the normal and the abnormal in paramediation is not reducible to overfitting and underfitting. This is demonstrated by the paraesthetics exhibited by the case studies. Even when these may not be quantitatively apprehensible as overfitting and underfitting, they may be qualitatively apprehensible via the aesthetics of the algorithmic paranormal: artifacts, uncanny repetitions, and disjunctures. Paraesthetics are more closely related to the qualitative aesthetics of glitches than to the quantitative statistics of over- and underfitting. It is by examining these apparent artifacts and aesthetics, as well as how and why they manifest to human perception, that we can learn something about the way that an underlying system is attempting to configure appearances – much like studying glitches (see for example Menkman, 2011).

Egrecorp was one attempt to do this. As discussed, through the game design process, I attempted to glitch familiar design conventions to interrupt their seamless coherence and demonstrate their construction. However, this proved to be a challenge, since interrupting this coherence simply made the game more difficult to play. Unlike films, where cinematographers can play with or interrupt familiar cinematic conventions to reveal their technical means of articulation without compromising the act of viewership, games require active participation from their audience, and thus require some level of coherence. This demonstrated a tradeoff between using paramediation to create a compelling, seamless, and immersive appearances on the one hand, and interrupting the experience of these appearances to show paramediation explicitly on the other.

However, I hypothesize that certain types of algorithms would be more perceptible as such; namely those that arrange phenomena to correlate to user activity in a *less linear* manner. That is, rather than responding to user interactions in a discrete and predictable manner, the most salient influence of paramediation would manifest through algorithms that arrange appearances in a way that is less overtly predictable. In fact, I attempted to implement algorithms of this kind by creating voting simulations in an early version of the game, which involved parameters that would change in complex ways according to player actions, and even on their own. But these simulations were too complicated to reveal any clear correlation to user activity. This demonstrates, once again, a balance between the normal and the abnormal in algorithm design: designing algorithms that respond directly to user behaviors in a consistent, linear, and normal way may recede from apprehension entirely, but those that have more sophisticated and spontaneous behaviors may appear to be unrelated to any observable norms.

Here the earlier theoretical distinction between interactions (between users and algorithms) and para-actions (between algorithms and other algorithms) is useful: it is the latter that manifest more unpredictable and spontaneous phenomena to users. Egrecorp involves the design of both user interactions and algorithmic para-actions, the latter of which articulate 3D worlds and environments where virtual objects can interact with one another. The relationships between para-actions are *internal* to the game system, rather than found in the ways that they interact with user behavior. To actually make paramediation perceptible is a challenge that requires a fine-tuning of the relationship between interactions and para-actions. It has little to do with exposing these actions and their relationships alone.

7.1.3 Playing Out

One aspect of this raised by both the case studies and the interviews was the role of procedural generation or *playing out* in the design of algorithmic rhetoric. As Francis put it,

> *... the fact that maybe two people could mutually agree on maybe how a system is structured, how it works, but have different ideas of what the implications of that structure was, meant that if you build a simulation according to those rules, and you showed a different outcome than what they expected, that would be more rhetorically powerful in some way ...*

The "rhetorical power" here is that letting rules *play out* from constraints communicates something about the implications of these rules (insofar as this produces consistent results without obvious human interference, as Francis notes). Palantir Gotham operates in just this way, by letting rules about associations among crime data *play out* from this data as an input. The "rhetorical power" of the system is enabled by the procedural generation of crime connections via input data and modifiable rules.

In turn, the analysis of Dwarf Fortress first raised the relationship between simulations, games, and emergence, which was later addressed in the interviews. It demonstrated the relationship between codified rules and emergent phenomena, and how this relationship becomes an object of design in its own right that involves an attention to human perception. For both simulations and games, codified rules are configured to produce emergent phenomena. The difference is that simulation rules must correspond to given norms, and then must be configured to produce results that correspond to other given norms, whereas game rules are designed more flexibly to generate whatever result that a designer wants to "show" (Francis) to an audience. A simulation is concerned with the correspondence of both algorithmic rules and results to the given, whereas a game is concerned with phenomenal results: their congruence to the given on an aesthetic register (if at all), irrespective of the realistic appearance of rules.

Unlike games, the credibility of a simulation depends on a scientific epistemology of verification and validation, where rules should correspond to results *consistently* and without obvious human *interference*. As Francis describes, a correspondence between rules and results that is consistent and independent from human intervention indicates causality. The hope is that, if both rules and results in simulations correspond to given phenomena or norms, then their correspondences indicate causal processes that could be accepted as true. On the other hand, the rhetorical power of the relationship between codified rules and their emergent results may be compromised if there is a perception of too much determination of the given by human interference. The challenge is that human interference with the given is inevitable, insofar as data and rules are designed in the first instance to make certain results appear, with certain consistencies or emphases. The legitimacy of phenomenal results depends on the supposed lack of interference in the design of rules for manipulating the given. This is a tension at play in

Palantir Gotham, which is staked on deriving causal relationships from data, but which also admits significant human interference – a tension which we can characterize once again as balancing the normal with the abnormal.

If games concern story generation through procedural play, we might say that simulations concern story telling through procedural proof. While the former is concerned with generating mostly original stories that subjects feel they are producing on their own, the latter concerns the capacity to articulate a relatively stable or consistent pattern of events – something comparatively more normal than abnormal. At the same time, the interviews with artists and game designers show how algorithmic phenomena may be designed to cohere with given norms in a qualitative or aesthetic way. Francis discusses the ways that algorithmic rules can be simplified or made more elaborate to enhance their credibility and accessibility: you can "round off the edges to reach a larger audience," and "more people will hopefully kind of see the point that you're making."

Here *interaction* can be involved in configuring the phenomenal effects of algorithm operations, as in Palantir Gotham: the credibility of crime narratives emerges through a process of interacting with the software's algorithms, rather than using algorithms to produce these narratives automatically. This is something reflected in the interviews as well, where for Francis:

> *... I suspect that the act of playing and interacting, often you're embodying a character, and you're kind of driving the decisions just puts you in a state that's maybe more receptive to seeing things through a different perspective, or thinking about things differently.*

Although Francis associates the rhetorical power of procedures with "embodying a character," we might also consider how the agency afforded by playing and interacting with rules enhances the credibility or effectiveness of what algorithms are designed to make appear.

7.1.4 The Trap of Design

Enhancing the effectiveness of paramediation, such as through interaction or immersion, is linked to the idea of deception and pacification. This idea first appears in what Oscar calls the "trap" of algorithmic experience design: the task to improve algorithmic experience may only optimize the imperceptibility of certain algorithmic harms. Put in the terms of paramediation, enhancing the effectiveness of paramediation depends on minimizing perceptions of a 'poor algorithmic experience,' or avoiding the rupturing aesthetics of the algorithmic paranormal. This was evidenced in the case of TikTok: the task to improve algorithmic content recommendations and minimize their harms could distract from their broader capacity to optimize particular behaviors overall. This is to say that the most significant, influential, and pervasive effects of algorithms have been designed to be innocuous, in contrast to the blatant harms of algorithmic misfires. Acknowledging this, a theory of paramediation encourages designers to acknowledge how the task to improve algorithmic experiences may only to be to optimize their power in other ways – namely to enhance the power of paramediation.

But for designers to acknowledge this, they must confront an impasse in design. While design work recognizes that design can deceive people by arranging appearances for them, its main tools for confronting this fact are to critique these deceptions either on the *ethical* basis of a good experience (e.g., whether the designed appearances feel good) or on the *logical* basis of a contradiction with algorithmic operations (e.g., whether the designed appearances correspond to actual algorithm operations). While such critiques might apply to evaluating textual statements, they are frustrated by the operations of paramedia that are designed to adapt to how people perceive them. Thus ironically, when Oscar distinguishes his design practice as addressing the experience of "people" rather than trying to determine the experience of "users," the issue is that paramediation actually does the former already; that is, calibrating experience not through direct control and regulation but sensitive, adaptable, and indeed qualitative responsiveness to human perception. The tension this raises is that, by implementing designs that address experience rather than determining it, we are not necessarily free from algorithmic harms; we may simply find ourselves in the business of improving the effects of paramediation.

Designers and explainable AI researchers find themselves trapped in this tension: they must eliminate algorithmic harms, but without eliminating algorithms entirely. They must explain precisely what it is that algorithms do to users, but without preventing the capacity of algorithms to operate in the first place. They must depend on the capacity of users to raise concerns about algorithms, when in fact their goal to improve system operations amounts to diminishing this very capacity. Issues arise when the proper functioning of algorithms is closely related to its harms, and can not be remedied with good design. To this end, the most viable solution may be to throw the baby out with the bathwater, as Hendrik remarks in the case of the COMPAS algorithm.

But as Hendrik also observes, the problem goes much deeper, since deception in algorithm design can apply just as much to developers as it does to users:

> *I was actually in a lab on deep learning ... And there's a bunch of people, very clever kids, and they are like, working, for instance, on tracking algorithms, right. So they spent like three or four years building a super good algorithm, looking at a video where they try to track Yoda in Star Wars. They've never even thought about what these tracking algorithms in the end will be ... And these are PhD level kids, and that's very dangerous and scary. But it's also probably not a coincidence that there's like a difference between how they are trained. And I think, yeah, we should definitely work against that. And help them understand the repercussions of the work.*

Here, the task for designers is to acknowledge patterns of deception and make them apprehensible to others, but this may ultimately be an extracurricular practice beyond their assigned work. While designers working in AX or explainable AI recognize that algorithmic systems can exploit users and frustrate positive experiences of them, their position of work makes it difficult to reveal these issues without compromising the functioning of the system. As Hendrik jokes, "... if you add explainability to the mix, you're kind of ruining all the fun in a way – you're kinda like ruining the purpose." *The purpose that is ruined here is a carefully crafted paramediation.*

7.1.5 The Truth of Paramedia: Spectacle or Mysticism?

Inversely, paramediation challenges organizers to acknowledge how thoroughly algorithmic systems arrange appearances beyond mere spectacle. As said, Hamid's references to *spectacle* follow roughly from Guy Debord's theory (Debord, 2012): both stress how constructed appearances can obscure conflicts and asymmetries on another material register, effectively reducing critical considerations to reforms that fail confront this suppressed material reality. Debord in particular stresses the partial nature of these appearances and the way that they advertise their own comprehensiveness nonetheless. But while algorithms can certainly be seen to participate in the construction of an apparent totality that obscures material conditions, this is not all that they do – and far from what they are designed to do. Instead, we would do well to identify how algorithms arrange new appearances alongside given ones. These appearances do not feign comprehensiveness and they do not stand in for the given. It is rather as if they are mixed or integrated in with the given via a logic of correlation, to formulate phenomena that inherit its characteristics, intensifying them or composing them in a new arrangement.

This is something that the designers know well: the arrangement of appearances is concerned with the realization of a particular task (e.g., optimizing engagement), rather than representing or obfuscating the given. To be sure, the organizers interviewed for this study are also sensitive to this function of algorithms, reflected namely in a concern with decontextualization. But their work would be strengthened by declaring explicitly that algorithms do not simply misrepresent the given in a spectacular way, but articulate data collected from it in a way that manifests particular appearances, which have particular functions, with an ambiguous truth value. This is to acknowledge the limits of an epistemological critique of incomplete data algorithmic bias, which supposes to bring the representations of algorithms more in line with the given, rather than acknowledging their capacity to manifest an "alternate reality" (Jamie) that is partially derived from the given and partially constructed to deviate from it.

Altogether, the case studies problematized the epistemic status of the appearances arranged by algorithms, which was an issue reflected in the interviews as well. These appearances are neither entirely true nor false, but derived from the given and then designed to deviate from them in a controlled manner, according to qualitative and quantitative assessments: synthetic images in I2I are derived from source images, events that evoke causality in TikTok are derived from data about actual user interactions, and crime narratives in Palantir Gotham are derived from crime data – and yet, the function of each algorithmic system is to articulate appearances that deviate from this given data in some way. This is the same function that Jamie described as decontextualization: a narrative is derived from data about someone, which has been "decontextualized" and subsequently "re-aggregated" in a new way. This construction of an "alternate reality" through data is not entirely true nor false, but partially true and thoroughly designed.

This formulation recalls Deleuze and Guattari's philosophical oeuvre, which insists on the "mutational" character of concepts – their transference, cross-pollination, and appropriation from one system of thought to another – as well as their contingency on the discursive and philosophical systems that they comprise (Deleuze and Guattari, 1994). For them, it is less insightful to identify how much a concept corresponds to an external reality, than it is to examine how a concept participates in a specific system of conceptual relations that lends it consistency and significance. The point here is not to undermine any consideration of the materiality or provenance of concepts, but to acknowledge the widely diverse regimes of perceptibility and conceptualization that can be extrapolated from it (Deleuze, 1988) – and this requires that this regime possess a consistency or coherence with respect to itself, even when it does not reflect reality absolutely. Decontextualization and recontextualization (or deterritorialization and reterritorialization in Deleuze and Guattari's parlance) is similarly less concerned with an allegiance to the truth than with its ability to derive something from existing circumstances that can be re-arranged to establish a relative coherence which becomes operative in a new way.

This understanding may be able to resolve a tension expressed by the artists in the interviews: that there appears to be a tradeoff between making visible algorithm materiality and making visible its emergent effects. Show too much of an algorithm's internal operations, and you can lose sight of its significance in the world; focus too much on an algorithm's effects and risk sliding into "mysticism." But perhaps this dichotomy is trapped in the notion of a phenomenal "spectacle" that occludes the material truth. What is also at stake is how exactly algorithmic phenomena are designed to have particular effects, namely by arranging appearances according to what is regarded as given, rather than according to the artificial determinations of algorithm logic alone. A theory of paramediation attends to how phenomena are arranged according to various considerations about subjective perceptions, algorithmic constraints, and design goals – which are for their part informed by other social factors. Instead of seeing the role of design to "make visible" either algorithm materiality or aesthetics, we might acknowledge how algorithm design is concerned with the very relationship between algorithmic rules and phenomenal results. This special kind of design, the design of paramediation, takes neither algorithms nor algorithm logic as its sole object, but namely the relationship between this logic and its phenomenal arrangements. Could it be 'made visible', like algorithm logic?

This type of design was certainly made visible through the process of designing Egrecorp, while perhaps not through the game itself. As the game designer, I had to develop game sketches and to play the game to get a sense of how it worked and might be perceived. When I received initial feedback from playtesting, I continued to play the game myself to identify whether it corresponded to feedback from playtesters. Although I was programming and working with code, I was namely designing the phenomena that would be evoked by this code.

The code was a proxy for treating these phenomena, as well as their ultimate perception by playtesters, as an object of design. And yet, it was difficult to demonstrate this fact through the actual game. While I tried to display the design of the relationship between algorithm logic and phenomenal arrangements through the game mechanics of the Debugger, playtesters did not comment on this relationship explicitly. Perhaps, the most effective demonstration of paramediation would be one that actually invites playtesters to create their own game. Insofar as I am the one making the game for them to use, their perception of paramediation can only go so far. But the process of constructing paramediation reveals it.

This suggests that another way to evaluate Egrecorp would have been to tell playtesters about the theory of paramediation *before* they played the game. While I planned the study to evaluate whether the game could demonstrate a theory of paramediation on its own, it may have been more effective and generative to identify how knowledge about the theory of paramediation would inform interpretations of the game. Instead of trying to evaluate whether the theory could be made perceptible through an algorithmic system, this would be to evaluate what the theory itself makes perceptible about an algorithmic system.

7.2 Limitations

Paramediation is a process afforded by algorithmic techniques, but which is nonetheless inseparable from a *design process*. I contend that the process of designing paramediation is an aspect of algorithm design: beyond designing algorithms to perform more effectively – by attending to their procedural complexity or statistical results – they are also designed to appear in particular ways by attending to the phenomena that they manifest. One limitation of this study is that it did not interrogate instances of algorithm design that have little bearing on the perceptible. Based on the results of this study, it is only possible to conclude that designing appearances is an integral aspect of some algorithm design, though perhaps not all. To what extent is it an aspect of all?

To answer this question, future work would need to consider algorithm developers (or algorithms) that do not explicitly address the social consequences or appearances of algorithmic operations. From my own experience in applied algorithm development, I hypothesize that paramediation is an important aspect of this work as well.

Another, more theoretical limitation of the study is that while paramediation is a way of describing an aspect of what algorithms do and are designed to do, there is not a null hypothesis that could refute the theory. This is a limitation of most critical scholarship involving algorithms: theories are put forward that do not involve any criteria for refutation, instead raising innumerable concerns about properties, aesthetics, and consequences of algorithms. Indeed these theories are ways of framing algorithmic systems, in order to perceive, conceptualize, and discuss them in new ways. They are not totally distinct from "folk theories" of algorithms (Eslami et al., 2015), except for the fact that they involve an attention to algorithm implementation details by experts.

The contribution of a theory of paramediation, then, is to focus attention on a particular aspect of algorithm operations and effects, as well as to center how this aspect is perceived and designed to be perceived by different people. Nonetheless, while the theory is evidenced by the study, it seems intractable to assert a widely generalizable and normative claim about how algorithms operate or are designed. A theory of paramediation should therefore be considered with respect to particular algorithms and algorithmic practices on a case-by-case basis. It should push us to acknowledge that algorithms can be designed in a wide variety of ways according to an immense variety of human practices, far beyond simply regulating activity according to rules.

A final limitation of the study – or perhaps a contradiction – was that while the goals for game design were not narrowly defined in order to experiment with and generate implications of paramediation, the game was nonetheless tasked with demonstrating these implications to users. The task to do both at once – to generate implications of paramedia and to explain them – did not lead to very effective results. Future work might dedicate more focused attention to either one of these tasks, but should take care not to conflate them. As well, playtesting the game could have been used to identify what the theory of paramediation makes perceptible about the game, rather than whether the game makes the theory understandable.

7.3 Implications

In this final section, I return to some of the theoretical issues raised in the development of a theory of paramediation, discuss them in light of the study results, and address the implications of a theory of paramediation for algorithm design and criticism.

7.3.1 Beyond Computational Ontology

A focus on the ontologies of algorithms cements a dependency on paradigms of governance, subjectivity, and media criticism that treat the impartiality, transparency, and flexibility of decisions as a limit to oppression. From this view, the more impartial, transparent, and adaptable computational decisions are, the less suspect they are. With respect to TikTok and Palantir Gotham, this is to ensure better representation of minoritized populations in data, or to exclude data about race from certain analyses. Meanwhile, for I2I, the transparency of algorithm operations is held to prevent their phenomenal harms. But from the vantage of paramediation, these responses fail to address the capacity of algorithmic systems to arrange phenomena to certain ends: to direct attention, to produce a coherent crime narrative, to produce an evocative image.

While epistemological criticisms of algorithm ontologies can contend with problems of data bias, algorithmic bias, and transparency, we see how these criticisms motivate designers to respond to these criticisms rather than their effects. This can motivate, for instance, systems that are designed with mechanisms for ensuring algorithmic transparency in a way that alleviates concerns, like a "placebo" (Vaccaro et al., 2018), or those that purport to avoid "discrimination" by excluding explicit racial categories from analysis (Bratton, 2018) or simply including more data (Bratton, 2022). Consequently, responses to algorithmic harms that conceptualize them in terms of regulation only *displace the power of algorithms from regulation to paramediation*, which raises new – and less salient – problems for social justice. These new problems are inapprehensible to traditional paradigms of algorithm criticism insofar as they treat appearances of algorithms as evidence of underlying ontologies, decision-making logics, regulative norms, and their opacity, rather then consequential and strategically arranged phenomena in their own right.

A theory of paramediation acknowledges that algorithmic systems are designed to evoke particular perceptions and actions by arranging phenomena, and this can involve designing systems in such a way that certain criticisms no longer apply to them – such that people view their operations as harmless in the case of TikTok, or coherent in the case of Palantir Gotham. This challenges theories of technology design, such as Bruno Latour's actor-network theory (Latour, 2007) and Andrew Feenberg's theory of technosystems (Feenberg, 2017), where the harms of technology and their perceptibility are contingent on their inability to represent the interests of users or stakeholders. These theories fail to account for operations of technology which are not concerned with meeting user needs and addressing user concerns, but with evoking perceptions such that these needs and concerns are satisfied by other means. Latour and Feenberg, following theoretical paradigms from science and technology studies (STS), argue that technology design will tend to converge on a metastability that meets the interests of stakeholders, or else it will be contested and fall into disuse. This presupposition of ontological harmony between stakeholders downplays technical strategies designed to influence perception and satisfy desires without upsetting or contradicting them.

A theory of paramediation thus challenges the notion that the social effects of algorithms are reducible to matters of epistemology and representation. An epistemological approach to analyzing algorithm ontologies evaluates whether algorithmic representations objectively correlate to reality in fact, or otherwise to some equitable ideal. This approach implies a particular ethical imperative: algorithms should implement an ontology that adequately represents the world without misrepresenting it, and this correlation can be enforced to bring a more equitable world into fruition. It follows that, for every aspect of society – from healthcare to law enforcement to entertainment – a hypothetical algorithm exists that could respond to that aspect in an ethical way that does not cause harm.

This view is reflected in arguments that, to design more equitable, less harmful algorithms, we need more data (Beller, 2021), or even a better system for valuing data that incentivizes data collection of the right kind – against the system of value manifested in capitalism (Joque, 2022). In either case, the idea is that the data and algorithms we have today are incapable of representing individuals comprehensively and fairly, either because the scope of their analysis is incomplete and therefore partial, or because it is driven by the wrong values, and is therefore partial by design. Following this epistemological critique, we could either collect more data, to saturate the partial data with more robust alternatives, or change the system of values that drives partial data collection and analysis in the first place. Notably, these appeals to algorithm reform mirror earlier ones by specialists in computer science (for example Datta et al., 2016) – only now they center the influence of capitalism rather than human prejudice or error. We are not told concretely how a new system of values or ontology would enable us to implement more equitable algorithms; we are only told that the data and ontologies we have now are not enough.

7.3.2 Beyond Representation

Paramediation connotes phenomena that have been thoroughly constructed – artificial arrangements of phenomena. However, despite their being constructed, paramedia cannot be definitively classified as untruthful since they operate *alongside* and according the given, or that which is already accepted as truthful in data. Key to a theory of paramediation is that algorithmic operations cannot be evaluated simply by whether they correspond to an existing truth, insofar as their goal is to produce new phenomena and evoke new perceptions that correspond to the given in the first place. From I2I to TikTok to Palantir Gotham, these phenomena may be derived from data that is unambiguously accepted as given, but arranged in such a way that the significance of what is given is changed. This, beyond merely misrepresenting it. This raises a unique problem for criticism. While legal regulations or statistical norms can be disputed if they fail to represent material conditions, paramediation points to algorithmic operations that intervene in the given by making certain things about it perceptible – not representing or misrepresenting it outright.

Paramediation participates in what Lyotard called paralogy: the construction of meanings that are subsequently evaluated on account of their practical efficacy rather than their epistemic accuracy or generalizability (Lyotard, 1984). Algorithms arrange phenomena in relation to data collected about various events and phenomena, and then optimize these arrangements according to certain heuristics – engagement, believability, positive experience, etc. The arrangements that result, or paramedia, do not contradict the given nor adhere to it absolutely; it is rather as if they are mixed among, conjured with, or composed from it. Whether this composition is designed to be harmonious – evoking seamless virtual causalities between phenomena like in a musical score – or disjointedly revealing inconsistencies and interruptions between otherwise harmonious phenomena – depends on the task at hand. In either case, algorithms can calibrate and optimize the composition of phenomena to manifest a *coherent regime of sense*, by operationalizing data about their material conditions and subjective reception.

Indeed, there is growing talk about the "fictional worlds" or "imaginaries" that algorithms can manifest (Bucher, 2017). The speculative crime narratives synthesized by Palantir Gotham, in advance of their validation, might be some of them. But what is the epistemic status of these algorithmic fictions? If they are regarded as misrepresentations, they are deceptions that preclude accurate or holistic perceptions of reality, inviting speculation about an alternative algorithmic mediation that would be more comprehensive, less fictional. This elides the fact that algorithms have agency precisely because they construe the given in a new way. As the case studies in this study illustrate, the design of each algorithmic system involves decisions about what to sample, how to articulate it, and what heuristics should inform the revision of these parameters. For paramediation, these parameters are designed with an attention to what algorithms will make perceptible, in order to orient perceptions toward particular ends. Algorithm design is not so much about representing the given accurately or even comprehensively, but ensuring that its mediation meets a particular end.

A key significance of this is that, in many circumstances, an analysis of algorithm logic cannot have full purchase on the consequences of algorithms. Algorithm logic is first designed, and then modified according to data, to evoke particular appearances and have particular effects, which involves a regard for their subjective reception. On the one hand, this should encourage us to identify how designing for human perception is an aspect of algorithm design broadly conceived. On the other, it holds that certain aspects of algorithm operations are inapprehensible without a regard for how they appear, and namely, how they are designed to appear. This demands attention to the voices of those who address the consequences of algorithms in terms of their lived experiences with them, rather than in the language of technical expertise. This is consistent with Ned Rossiter's proposal for "paranoia as method" (Rossiter, 2017): the recognition that even though people may not know exactly how or why algorithms affect them, their theories about these effects can be more effective at grasping their consequences than an informed regard for their code. As Oscar put it, "They don't have a name for it. But they suffer."

Rather than representing or misrepresenting the given (mimesis, where what is important about a mobile device notification is whether it reflects the given: a message has been received), paramediation coordinates appearances according to the given to make something new perceptible. And rather than simply augmenting human cognition with the ability to apprehend more aspects of the given (prosthesis, where what is important about a mobile device notification is that algorithms enhance our ability to perceive the given: a message has been received), the notion of paramedia emphasizes that these mediations of the given are thoroughly constructed, often in ways that do not emancipate human activity. Paramedia articulate phenomena in ways that do not reflect the given absolutely, but nonetheless correlate to it in some way. A theory of paramediation thus motivates us to move away from the notion that algorithmic activities should be evaluated according to whether or not they correlate to the given: paramedia always does; the question is how.

To this end, the notion of the paranormal may, rather than 'sliding into mysticism,' help explain the epistemic stakes of paramedia: these are phenomena which may not have an origin that can be decisively localized in algorithm logic, data, and effects, but nonetheless emerge through the correlation of these factors, and have clear consequences for subjects who perceive them nonetheless. Paramedia are neither absolutely true nor false: while they manifest to perception conspicuously and are derived from existing circumstances, their epistemic claims are indecisive. Paramedia may appear as symptoms of processes that are not directly apprehensible, as if they emanate from unknown conditions or laws. And these symptomatic appearances are irreducible to algorithm logic, insofar as they are designed to make something perceptible that is not written explicitly in the given code: a sense of causality, a threat, a premonition. Acknowledging such paraesthetics can be useful for attending to algorithmic techniques, mechanisms, and practices designed to influence perception.

7.3.3 Plurality of Sensemaking and Coherence of Sense

If not the epistemic validity of algorithm ontologies and representations, then, how should design or criticism address the social consequences of algorithms in practice? First, at minimum, paramediation encourages us to acknowledge that algorithms have agency by enacting a deviation from the given: rather than representing the given, algorithms always intervene in its consequences and reception. In turn, this means that we must acknowledge the range of possible deviations that paramediation is able to derive from the given. Even if what is given in data were to be regarded as unequivocally true, the possible effects that paramediation can enact by correlating appearances to this data are considerably vast. This exhibits a *plurality of sensemaking* at the heart of algorithm design: a capacity to make sense of the given in a wide variety of ways.

Before paramediation is applied in practice, this plurality is bounded only by the capacity of paramediation to correlate to what is given. Without deriving from the given – without an input – paramediation is not at play. However, the application of paramediation *in practice* tends to constrain the possibility space of sensemaking according to a particular task – that is, parametrically – to I2I, to optimizing engagement, to producing coherent crime narratives. This tends to minimize the plurality of sensemaking afforded by algorithmic media by maximizing a coherence of sense, where the object of paramediation is to optimize phenomena such that they realize a particular effect. The process of sensemaking is thus constrained according to more rigorous parameters. We could say that it has been automated to realize a particular task.

This, however, does not necessarily need to be the case. While I2I, TikTok, and Palantir Gotham are chiefly concerned with realizing the coherence of sense, this is not all that algorithmic media can do. Could algorithmic media be applied, instead, to realizing a plurality of sensemaking processes, where multiple avenues for deriving and deviating from the given are opened up?

7.3.4 The Politics of Algorithmic Experience

This plurality of sensemaking is raised by the stories cultivated and disseminated by the organizers who were interviewed for this study: stories of human experiences with algorithmic systems that contest the narratives that these systems put forward. This *agonistic* encounter between subjective stories and paramedia challenges the notion that algorithms could be made harmless by designing them the right way; at bottom there is still a decision being made about how sensemaking should occur, on behalf of others. Subjective experience will always have a role to play in apprehending not only the phenomena that algorithms arrange, but also, as the interviews demonstrate, for whom and by whom these phenomena are arranged.

For the designers interviewed, an attention to lived experience motivates a demand to design algorithmic systems from the bottom up, so that they support user needs and capacities, rather than determining them. Critically, this imperative to grasp the design of algorithms in terms of experience is less concerned with access to, authority, and control over algorithm operations per se, than with the experience that these operations manifest. This focus is reflected in the work of the organizers, game designers, and artists as well. For Neilson, it is a matter of "reclaiming" the lived experiences of proprietary game algorithms, while for Jamie and Hamid it is practiced through community storytelling. Here the politics of these algorithmic experiences have less to do with precisely how they are implemented than whether they afford a plurality of sensemaking processes, make them more accessible, or mitigate them.

A theory of paramediation accounts for these politics of algorithmic experience, their process of being configured by design, and their effects. For Oscar and Hendrik, users attest to what they experience when they use algorithms, which is then used to inform another design. For Hamid and Jamie, the subjects of algorithmic systems attest to their experiences through stories, as a way to challenge the seamless appearances of information-driven policing algorithms. And for Neilson, Grayson, Francis, and Danae, what is interesting about algorithmic media is their capacity to inspire a kind of indeterminacy which becomes apprehensible by considering what is perceptible during play.

Does this indeterminacy confer plurality of sensemaking? Perhaps not. Relatively indeterminate interactive mechanisms, like in Palantir Gotham, admit a kind of ontological plurality that nonetheless enhances the coherence of a final perceptible result. This suggests that enhancing interaction and participation in algorithmic systems is insufficient for supporting a plurality of sensemaking. As well, it suggests that a plurality of sensemaking is not supported by simply collecting more data – all this data can still be mobilized in the service of manifesting a coherent sense. While Palantir Gotham solicits user participation specifically in order to achieve the ultimate end of constructing coherent crime narratives about others, facilitating a delimited form of participation in the decisions of the TikTok algorithm, to give another example, only supports its goal to optimize attention and engagement.

In the end, if the people most affected by these systems were able to govern their sensemaking parameters without restrictions, the utility of these systems would be compromised outright. This suggests that restricted and hierarchical access to control over algorithms is – far from something that can be ameliorated by design or a new system of values – is inherent to many of their operations, and demands an agonistic politics that can constantly confront them.

Here we return to the problematic raised by algorithmic experience design: how do we challenge, let alone conceptualize, exploitation when it is explicitly designed to be perceived otherwise? While epistemological and Marxian approaches insist on exposing the material reality underlying the algorithmic spectacle, they may have to confront the fact that the scandals of the source code are not very secret after all: even when people know that their data is being collected, algorithmic experience can be designed such that they keep on using platforms nonetheless. And this is not simply because these platforms monopolize our means of communication, such that we have no other choice, but because they have been carefully designed to arrange coherent appearances that attract our attention and incentivize certain actions. Meanwhile, some theorists argue that this is not all that bad, turning Marxist theory on its head to argue that we need to design better systems of alienation, counterposing what is viewed as a fantasy of immediacy to a pragmatic attitude toward technical media design (Cuboniks, 2018). This addresses that technical media can be designed, but it does not address the politics of this design, and namely how this design can be imposed on one by another.

For paramediation, we first acknowledge the irrevocable power asymmetries implicated in algorithmic operations, insofar as they are designed by one to make sense for another. Putting a non-capitalist system of values or vanguard of workers in control of these sensemaking processes in no way absolves them of their power. There is a power in paramediation that is irreducible to who or what is represented by algorithms and whether these representations are profit-driven. We must acknowledge that what makes algorithms unique is their capacity to arrange appearances according to other factors automatically, which introduces a potent capacity to forestall plural and critical processes of making sense. Algorithms are by no means the first media to do this, but they have radically enhanced its effectiveness and dynamic flexibility.

7.3.5 Beyond Norms

If the politics of algorithmic experience are staked on an opposition between the plurality of sensemaking and particular coherencies of sense, this could be read to suggest that what we are dealing with, after all, is a *regulation* of sensemaking by algorithm logic. Is paramediation, in the end, a way of regulating our capacity to make sense, which should be critiqued on the basis of constraining sensemaking too narrowly? In that case, would not the goal for design and critique be, ambiguously, to *deregulate* sensemaking?

These questions can be addressed by attending to theories of normalization, or the use of statistical averages about populations to classify and govern individual activities. For theorists of governmentality, human conduct can be managed by forcing or pressuring it to correspond to a norm, according to which behaviors can be compared, classified, and managed – this is the basis of *subjection*. Algorithms are well-known for establishing norms that are not rigid and fixed, but flexible, able to change depending on particular people and circumstances. Jürgen Link distinguishes between these types of norms as "protonormalization" and "flexible normalism" (Hall & Link, 2004). He argues that the former involves an "other-direction" to which individuals are subjected to – a regulation imposed upon individuals to treat them according to certain standardized criteria, while the latter entails an "inner-direction" by which subjects normalize themselves – the basis of *subjectivation*.

Challenging a neat distinction, Tobias Matzner emphasizes that data-based surveillance involves forms of flexible normalism that nonetheless depend on other-directed controls (Matzner, 2017). Looking to Foucault, he recalls the distinction between *normation*, which begins from a given norm and classifies the normal and abnormal according to it, and *normalization*, which uses data about populations to deduce a norm according to which decisions can be made:

> *"we have a plotting of the normal and the abnormal, of different curves of normality, and the operation of normalization consists in establishing an interplay between these different distributions of normality." (Foucault, 2007)*

In contrast to a fixed norm, "distributions of normality" are derived from data and thereby respond to this data flexibly and dynamically. But for Matzner, this does not mean that they do away with fixed norms, as data-driven surveillance systems nonetheless use these calculations to subject individuals to normalized social categories, like criminal, in the last instance. Thus we are dealing with systems that identify norms flexibly through data, but enforce them in a more rigid fashion like through a code of law.

For scholars of governmentality, such enforcement can be conducted in two ways: through direct coercion – moderating physical or biological processes that determine or constrain activity directly, imprisoning bodies or blocking access to certain resources – or through the arrangement of social categories, ideas, practices, and appearances that dispose people to act in particular ways. Following Foucault, David Beer points out that algorithms can do both: they are capable of both direct *material* interventions and socially mediated *discursive* interventions (Beer, 2017). Other scholars adopt Guattari's framework to describe this difference as one between "machinic enslavement" by a-signifying semiotics and "social subjection" by signifying semiologies (for example Langlois, 2008). Altogether, these formulations of algorithmic agency acknowledge that algorithms can enforce both rigid and flexible norms, either through material interventions or discursively mediated ones.

These theories still remain largely beholden to the idea of norms designed to regulate activity: flexible norms, norms derived from data, and norms self-imposed through techniques of subjectivation, are themselves seen to be regulated by normalizing algorithmic rules and calculations. Indeed, such a conception of norms appears to be the only way that we can gain purchase on the social consequences of algorithms with some degree of reliability; without it, we are left in the territory of theorizing the indeterminate movement of subjectivation

A theory of paramediation may be a way to resolve this tension. Unlike protonormality, where individuals are regulated according to a fixed norm (subjection), and flexible normality, by which individuals normalize themselves (subjectivation), *paranormality* involves the arrangement of perceptible phenomena according to multiple norms operating in dynamic relation to one another, through a process of parametric design. With I2I, for instance, we see how norms derived from source images are deviated according to certain heuristics to optimize certain perceptible effects. While the resulting images are generated from statistical norms derived from source images, and subsequently optimized according to heuristics that normalize their effects, the resulting phenomena do not normalize human activity in any strict sense. Through paramediation, certain aspects of phenomena can be normalized to consistently realize certain effects, but these effects do not regulate human activity themselves. Arrangements of phenomena can dispose activity to various ends – influencing a particular response toward an image, for instance – without precisely regulating it.

This consideration is distinct from Deleuze's notion of "dividuals" encoded in diverse data streams that inform their treatment in dynamic ways (Deleuze, 2017). Deleuze's concept points to the fact that the technological apprehension of subjects does not need to leverage discrete social categories, and can instead disaggregate aspects of behavior into "molecular" or "a-semiotic" encodings. But it does not attempt to address how these encodings are used to subsequently influence behavior – through subjection, subjectivation, or otherwise. For its part, paramediation shifts attention from data collection and data processing to how data is used to affect subjects in the last instance. While normalization is always involved in some way at the level of data formatting and statistical analysis, the capacity of algorithms to arrange perceptible phenomena involves a different kind of influence that is not normalizing in the same way.

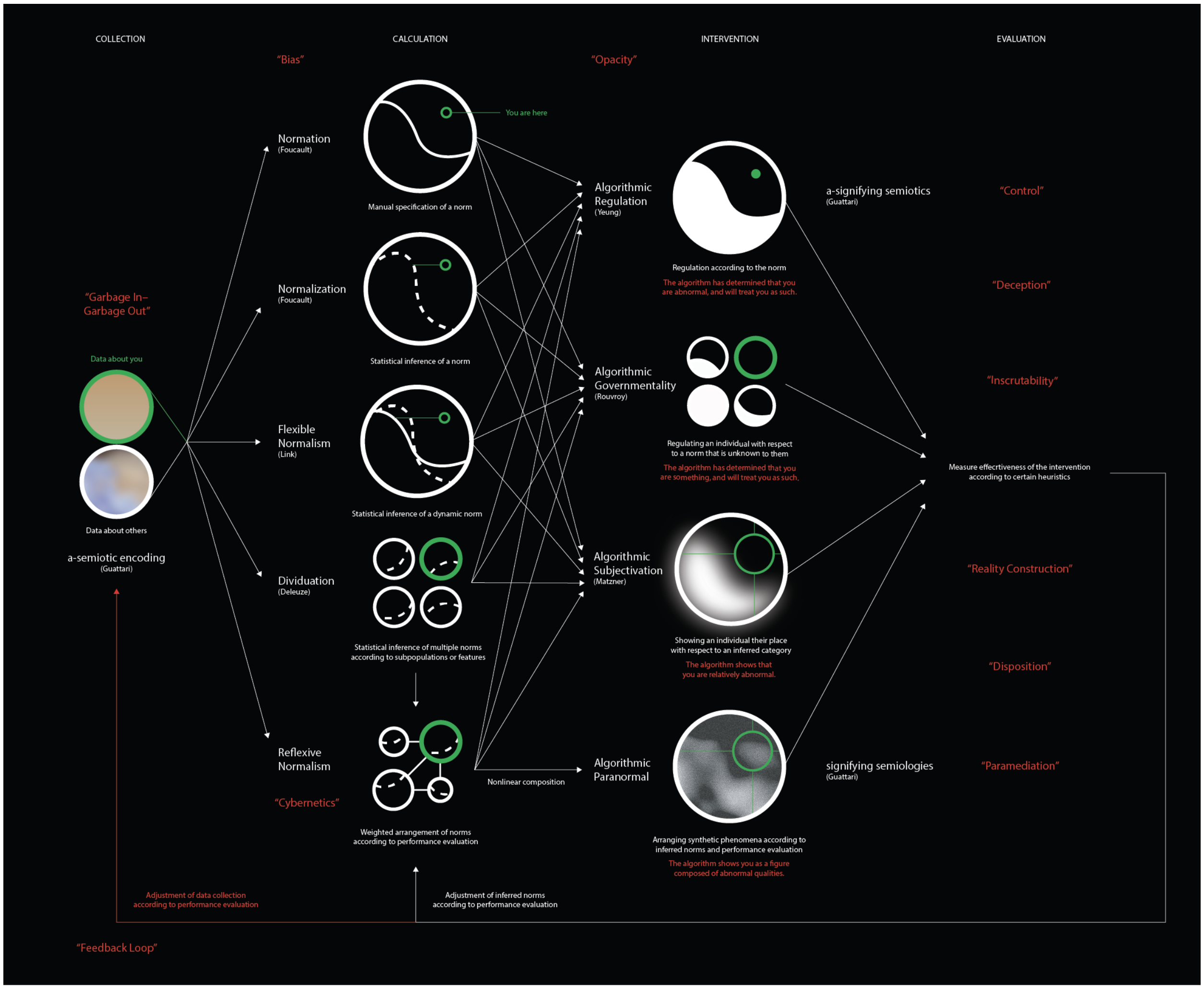

Figure 7.2: Diagram of different theories of normalization, as concerning data collection, analysis, and application. Paranormality (bottom path) is distinguished by using parametric design to produce perceptible phenomena that correlate to optimizing a human action, but without regulating human activity outright.

The designed coherence of sense through paramediation, then, is less a regulation of sensemaking than a particular way of orienting sense toward a kind of stability in measurable human activity. This coherence is derived from norms in data, deviated according to norms programmed into algorithmic rules, and optimized according to norms measured in how human activity responds to it. But the lump articulation of these norms is not a single regulatory norm, but a paranorm which involves dynamic correlations and affords interpretive flexibility.

7.3.6 Beyond Code as Law and Image

The notion of "code as law" (Lessig, 2000) makes sense if we conceive algorithms as rule sets that regulate activity, but falls apart when we acknowledge the perceptible phenomena that algorithms can arrange to a high degree of granularity, evoking perceptions without restricting activity outright. Even a "protocological" conception of algorithmic power (Galloway, 2004) – which attends to the way that decentralized and flexible rules can govern human activity – is based on the idea that computation exerts power fundamentally by regulating activity. And while some scholars argue that the distinction between algorithms and texts is that the former can "execute," enact, or perform their codes (ibid.), they remain beholden to the idea of code as a fundamentally regulative text like a code of law. According to this model, individual subjects are constrained by the regulations that media impose on them. It suffices to ensure that the regulations are equitable, or that legal policies can regulate them in turn.

Other, related models have been proposed. In particular, scholars for decades have developed productive analogies between algorithmic media and others, namely focusing on the relationship between computation and cinema. A cinematic conceptualization of media highlights the capacity of media design to arrange existing phenomena to evoke novel perceptions. Where Lazzarato brings the semiotics of cinema into dialogue with those of interfaces (Lazzarato, 2014), Anikina examines their implications in algorithmic art and videogame media (Anikina, 2020).

Deleuze's Cinema I and Cinema II books (Deleuze, 1986a; Deleuze, 1986b) also serve as a point of inspiration for algorithm scholars, because of the way that they taxonomize techniques for arranging images and sounds in space and time in order to elicit certain kinds of perceptions. Rather than a study of cinematographic techniques as such, the Cinema books aim to understand how percepts can be composed relationally in order to evoke certain sensations and thought processes. A similar investigation could be undertaken for paramediation, taking account of the capacity of such algorithmic processes as recursion, optimization, and procedural generation to arrange percepts.

Indeed, some scholars have pointed to the prospect of an algorithmic "Cinema 3.0" of networked, interactive, and algorithmic perceptions (Daly, 2010). The most concrete account is César Alberto Albarrán-Torres's analysis of gambling games (Albarrán-Torres, 2014), for which he theorizes the "procedure-image" following Deleuze's "movement-image." Albarrán-Torres invokes Ian Bogost's theory of procedural rhetoric (Bogost, 2007) to define procedure-images as "images that articulate interactive rhetoric." This accounts for their discursive function – their capacity to enact persuasive statements (e.g., "gambling is an adventure") – and their capacity to remediate or emulate other practices like playing poker.

Albarrán-Torres emphasizes the way that interaction and play is configured to yield narratives and discourses that shape subjects' perceptions of themselves. Notably, Albarrán-Torres's account of the procedure-image involves features of media that we might normally associate with textual or legal media, as opposed to the dynamic and recursive operations of algorithmic media. Namely, for Albarrán-Torres, the procedure-image requires direct physical interaction from players in order to operate, it enacts particular discourses from the persuasive to the colonial, and it is displayed in the same way to all players, regardless of who they are – in this sense procedure-images "do not discriminate" – much like a rigorous 'equality before the law.' Perhaps because of Albarrán-Torres's emphasis on "rhetoric," inherited from Bogost (2007), the theory of the procedure-image remains relatively indistinguishable from theories of textual media, which similarly aim to persuade and forward particular discourses, provided additional user input and interaction.

For other scholars, "Cinema 3.0" might be characterized instead by those algorithmic processes that elude direct perception – what the artist Trevor Paglen calls the "postvisual" in remembering Harun Farocki (Paglen, 2015), whose artwork responds to the capacity of algorithmic systems to sense phenomena and act according to them without direct human input. The idea of the postvisual is that algorithms can operationalize percepts as judgments without human (over)sight, raising concerns about the intelligibility and agency of algorithmic operations.

This notion can be misleading insofar as it implies that algorithms enact unprecedented forms of invisible coordination that other media do not. While algorithmic operations may not be immediately perceptible to humans, they almost always become this way – even if this is only realized after the fact, as a consequence of an algorithmic operation having been executed. Thus the postvisual is not the non-visual, but visualization by other means, and to other ends. While in extreme cases algorithmic decisions are enclosed almost absolutely inside a black box, every black box involves a tail of interactions and interventions that are perceptible nonetheless. Our task is to apprehend them where and how they occur, even when we don't have recourse to algorithm code or what algorithms 'see.' Indeed, Farocki's work is staked on rendering the postvisual to human perception, to demonstrate the implications of not seeing how algorithms see.

Effectively, the postvisual recapitulates the idea that algorithmic operations enact opaque regulations with their own agency, but it realizes a kind of counter-factual, as if these regulations could really be made visible, translated into images, as if they necessarily involved percepts and images. If 'code as law' is beholden to a textual conceptualization of algorithmic media, the postvisual remains beholden to a cinematic one, even while it insists on escaping from it. A more accurate model might be found in the *paravisual*: the displacement of perceptions to other means and ends, to different places and different times, which trouble our sense of being able to bear witness to everything that human subjects supposedly control. While technical processes that elude human apprehension have always been postvisual (or rather, previsual), what we witness with the paravisual are aspects of intentional, calculated *design* that elude perception. Algorithms participate in this design in a way that eludes perception, but this is not to say that they preclude it outright.

Paramediation is the use of algorithmic operations, which certainly elude human perception, to arrange, coordinate, and calibrate perceptible phenomena nonetheless. This involves elements of a cinematic conceptualization of computation, and of a postvisual one, but we should acknowledge the limits of cinematic view that inform these conceptualizations. From the cinematic vantage, what algorithms make perceptible should necessarily happen in a coherent time and place that is comprehensively subject to human authorship: when it does so, it is cinematic; when it fails to do so, it is postvisual. But paramediation involves a configuration of perceptibility that is not perceptible in itself, and which does not always have as its target something that could be conceptualized as an image, a film, a text, or even as a discrete media artifact. What has changed is not the expulsion of humans from sensemaking, but the troubling of our relationship to how sense is configured, operationalized, and distributed beyond the contours of discrete artifacts, explicit design choices, and subjects. What, then, should we forward as criteria for addressing the social implications of these media in design and critique?

7.3.7 Designing for Plurality in Sensemaking

This brings the political stakes of paramediation into view: when we acknowledge the capacity of algorithms to manifest a plurality of sensemaking processes, we are encouraged to resist the notion that their sole function is to apprehend the given accurately or direct it optimally. As said, such a notion motivates criticism to respond to the harms of algorithms by seeking to improve their capacity for representation, or otherwise to double down on the need for expertise that can improve algorithm ontologies. Accordingly for them, when we are confronted with human perceptions of algorithmic harms, these perceptions become mere indices of the *failure* of algorithms to represent the given adequately. But in fact, these perceptions are effects of algorithmic operations in their own right; they are not supposed to correlate only to the given, but also to novel arrangements of the given that manifest particular coherencies of sense.

If the goal is no longer to only improve how algorithms correlate to the given, we acknowledge their phenomenal effects that emerge from how this correlation is designed. The function of critique becomes to account for these possible correlations in all their plurality, identify the function of specific correlations in relation to specific social contexts, and question their restriction to unyielding coherencies of sense. Here, stories from lived experience, or practices for reclaiming algorithmic experience in other ways, both have a part to play. From this view, human perception and experience is no longer a mere diagnostic instrument for improving the correlation between algorithm logic and the given, but a target or subject of algorithm operations that can be raised to cast our understanding of algorithm sensemaking in a new light. A theory of paramediation demands that we raise these politics of algorithmic experience – who configures sense for whom and according to what criteria – to account for the capacity of algorithms to influence human activity that is irreducible to epistemology.

To this end, Matthew Fuller and Eyal Weizman's notion of an "investigative commons" may be instructive (Fuller & Weizman, 2021), where diverse perspectives and technologies are assembled together to apprehend circumstances from multiple overlapping points of view. Fuller and Weizman acknowledge that while technologies like algorithms augment our capacity for sensemaking, the configuration of sense remains a political act that should solicit participation from those who have a stake in deciding what makes sense. Building on the idea of an investigative commons, a theory of paramediation insists that the purpose of commoning sense-making and technological sensors is not only to apprehend the given in a more robust way, but also to account for the politics of configuring sensations. Paramediation shows that algorithms are not only sensemaking technologies, but also technologies for making sense in the interest of particular, plural, and political aims.

In this sense a theory of paramediation motivates us to bring Jacques Rancière's philosophy to bear on algorithm design: politics is not only a matter of configuring ontologies that order the world, but also of interrogating the phenomenologies that license this ordering (see for example Rancière, 2010). Acknowledging this, we must continue to find ways to balance an appreciation of the technical dimensions of algorithm operations, their norms and ontologies, with an understanding of what these operations are designed to make perceptible. The latter involves a capacity to influence behavior without subjecting it to norms and ontologies outright, and demands an attention to the lived experiences and perceptions that register this influence.

Altogether, this is not to argue that algorithms cannot be used to apprehend phenomena in the material world, to find statistical patterns in circumstances, to aggregate and discover facts that exist beyond the vantage of human perception. It is rather to insist that the capacity to play with these arrangements is plural and vast, and that the goal to simply improve what algorithms can represent tends to miss the point of what algorithms like those in I2I, TikTok, Gotham, and even Dwarf Fortress are designed to do in the first place. For designers, this might amount to designing systems to solicit more unstructured play, which is not optimized according to a specific heuristic. Or it could look like what Hendrik proposes for algorithm design: designing systems from the ground up in a way that is sensitive to how people interpret them – leaving space for their sensemaking practices alongside those of algorithmic systems.

These imperatives for design might be approached by acknowledging algorithms as agents in Wittgenstein's notion of language games, or dialogical interactions that inform the meaning of actions and speech (Wittgenstein, 2009). For Wittgenstein, the meaning of a term or a subjective understanding of expected conduct are elaborated through a process of interacting with its use over time. Such meanings and understandings thus imply communities, which take up certain kinds of interactions in common. But now, this configuration of sense through dialogue and interaction among subjects has become a target of technology design – this is the function of algorithms as concerns a theory of paramediation. Indeed, from the interviews, we saw how the algorithmic design of appearances, sensations, and meanings attains the capacity to form communities of sense.

Seeing that computer systems can structure meaning in this way, the philosophers Michael Hardt and Antonio Negri proposed that it paves the way to realizing a technologically-mediated commoning of sense (Hardt & Negri, 2000). For them, in Marxist fashion, the technology must only be seized in order to do so. But while Hardt and Negri reflect a Marxian desire for ontological harmony by seizing the means of perception and arranging them in an equitable way, an investigative commons (Fuller & Weizman, 2021) always demands an agonistic encounter, an engagement with forms of sense-making that may not even be acknowledged as such.

This is Rancière's gambit: a mode of sense-making inheres in every political regime at the exclusion of others, and the notion of including them all is not only to embrace an illusion of their total communicability, but also to adopt the pretense of every political order to posit its own sense-making as exhaustive. A theory of paramediation acknowledges how algorithms participate in the configuration of sense, which is not limited to collecting as many sensations as possible, so as to aggregate the most accurate and coherent sense. It also involves a power to make sense, which is a power that can be made more equitable only by subjecting it to the plurality of other perceptions, phenomenal arrangements, and designs. This is the task for algorithm design and critique. The project to alleviate algorithmic harms by improving their capacity for representation is to overlook the power to make sense outright.

Paramediation names the capacity to arrange phenomena, and thereby to influence perceptions, by coordinating them with data collected about activity or material circumstances. It marks a shift from *aesthesis*, or sense-making by human cognition and norms established in communities, to *paraesthesis*: the technical synthesis of appearances to correspond to given activities, appearances, or events with a high degree of precision and resolution. This encourages us to focus precisely on how algorithmic arrangements of phenomena are designed, and to what ends. It is not enough to say that they exist, and that they do not reflect the given absolutely. The question is always how.

Appendix

A) How GANs work

Before a GAN is trained, both the generator and discriminator have random parameters. The purpose of training is to update these random parameters so that they generate images and discriminate between images according to some criteria, instead of randomly. The discriminator neural network takes images as input data, in the form of an array of pixels, which each contain red, green, and blue color values – meaning that each image contains many thousands of values. These many thousands of values are input into many thousands of functions in the discriminator neural network, or "nodes," each of which contains a random parameter. Each node, depending on the difference between its input value and its random parameter, updates its random parameter slightly and produces a certain output value. These output values are then used as input for a following group of nodes, or "layer," and the same process repeats for multiple layers. Ultimately the values in the final layer are combined into a single function that outputs a single value. This value, typically a boolean (true or false) variable, indicates whether the neural network "labels" the input image as coming from the true data distribution or not.

The "error rate" of the discriminator is the number of times that it incorrectly labels target data as fake (false negative), or labels a generated candidate as real (false positive). Because it is initially random, the discriminator will begin with a high error rate: it will not successfully identify which images belong to the true data distribution and which are generated. When this error rate is high, the discriminator will randomize the parameters in its nodes more significantly. Over time, it will chance upon random parameters that label the input images more accurately, resulting in a lower error rate, and thus reducing the amount that it randomizes its parameters. Eventually this process should "converge" on a configuration of parameters that discriminates between true data and fake generated candidates with a relatively high accuracy. But the discriminator is only one half of the GAN.

While the goal of the discriminator is to decrease its error rate, the goal of the generator is to increase the error rate of the discriminator. It does so when it generates images that are difficult for the discriminator to distinguish from the true data. This is why a GAN is "adversarial": the generator is supposed to deceive the discriminator. Using the same process of conditional randomization as the discriminator, the generator randomizes its values the more that it fails to deceive the discriminator (that is, when the discriminator error rate is low). It thus tends toward parameters which generate images that successfully deceive the discriminator. These deceptive images bear a statistical likeness to those in the true data distribution, which typically appears to humans as a structural or aesthetic resemblance. This completes the process of training the GAN.

After training a GAN, image-to-image translation can occur when a single target image is used as input into the GAN, which serves as the starting point for the generator to produce an image. If during training, the discriminator converged on a configuration of parameters that is able to discriminate the true data distribution, and the generator converged on a configuration of parameters that maximized the discriminator's error rate, the images it generates will appear similar to the target data. However, this is not guaranteed to occur: even if a GAN converges during training, the generated outputs may be blurry, too similar to the input data, or altogether unable to execute an image-to-image translation task successfully. For this reason, GANs for image-to-image translation have been carefully designed to ensure that they converge on solutions which, irrespective of their statistical properties, appear to human subjects as if they solve image-to-image translation problems. This means that human perception is essential to designing and refining GAN models – to design a GANs is to design algorithmic display.

While GAN models can automate I2I without human oversight, human sight plays an important role in developing these models in the first instance. While GANs can be used for "unsupervised" learning, a type of machine learning that does not depend on human instruction to pair source data with target data, they nonetheless depended on supervision *during design*. The notion of unsupervised learning may be deceptive to nonspecialists: it does not imply that algorithms have arrived at particular solutions without human guidance; simply that they no longer require this guidance once they have been effectively designed.

For image-to-image translation, human supervision during design involves, namely, selecting an algorithm that is capable of displaying certain phenomena (e.g., choosing to implement a neural network), tuning its parameters to refine the phenomena it displays (e.g., designing a loss function), developing new configurations of algorithms that display these phenomena more consistently (e.g., generative adversarial networks), identifying input images that compromise the consistency of this display, and so on. These algorithmic and statistical innovations depend on a science of *computational phenomenology* to exist: a human practice of interpreting algorithmically synthesized images, identifying their artifacts, and developing computational techniques for altering them.

Despite the wide breadth of their applications, a given GAN model cannot solve every problem; specific models must be designed for specific tasks. Nonetheless, certain statistical techniques and algorithmic methods are generalizable across applications. Significantly, these generalizable techniques are not concerned with ontological properties of the images that they manipulate, but with the detectability and reproducibility of phenomenological features. Namely, they involve designing methods to produce outputs that are reproducibly coherent and generative; that is, consistently recognizable as images, but not because they mimic the target data absolutely (i.e., overfitting). Various approaches to coherence and generativity can be adapted across use cases, but may need to be modified to suit certain use cases or image types. Designing GAN models involves identifying the extent to which certain methods can be generalized to other applications, and to what effects.

To do so, phenomenal heuristics, or criteria for determining whether an algorithm solves an image-to-image translation task, are imperative. Phenomenal heuristics include identifying whether certain aesthetic, semiotic, or structural features transfer from the source image to the generated image, or from the target image to the generated image. The specification of phenomenal heuristics involves identifying phenomena in images manually, and then developing algorithmic techniques for identifying and reproducing them. This approach has resulted in, for example, the Learned Perceptual Image Patch Similarity (LPIPS) metric, which uses crowdsourced data from human workers on Amazon Mechanical Turk to train neural networks to calculate image similarity. Another approach called convolutional neural networks (CNN), inspired by neuronal behavior in the human visual cortex, uses a series of neural networks to scan small subsets of images to detect features like lines, curves, and shapes, which are then aggregated semi-randomly into larger neural networks.

That the process of designing I2I depends on phenomenal heuristics does not mean that it is absolutely dependent on human oversight – I2I involves methods for information processing that are not totally directed by human intervention. Nonetheless, these methods are designed, maintained, and tuned insofar as they can solve I2I tasks arbitrated by human perception: phenomenology ultimately determines the design and use of algorithmic techniques for I2I. While algorithmic techniques for synthesizing and detecting the realism of composite images can be generalized and adapted to other applications, human perception remains integral for evaluating the design of I2I algorithms. This fundamental principle is overlooked by descriptions of algorithms as systems that 'learn' how to recognize and reproduce images on their own terms.

B) Interview Questions

(1) artists

- Do you view algorithms as a subject matter in your work, as a medium, or as something else?
    - Which of your artworks demonstrate this?
- What elements of algorithms or algorithmic systems attract you in your work? Which do you tend to emphasize? Is this always intentional, or sometimes emergent?
- How do you think the experience of designing algorithms in an artistic setting differs from a non-artistic application?
- What are some particular challenges that come with making artwork with/about algorithms?
- What understanding of algorithms does art bring us?

and

- What do algorithmic glitches or artifacts convey?

(2) algorithmic experience designers

- What is a bad algorithm experience? A good one?
- What way of perceiving algorithms does user experience design enable us to take account of?
- Are there times that user experience design involves designing algorithms?
    - (if yes) Which of your projects demonstrate this?
    - (if yes) How do you think the considerations involved in designing algorithms for user experience design differ from designing algorithms in other contexts?
- How do the considerations of user experience design change when there are algorithms like machine learning algorithms involved?
- Does public attention to the consequences of algorithms influence what user experience design must be sensitive to?

- (if yes) In what contexts is this most salient?

(3) community educators

- What are typical barriers to understanding how algorithmic systems operate?
    - What is an example of an algorithmic system that you have worked with which demonstrates this?
    - (for each) Can it be overcome? How?
- What kind of techniques and resources do you try to develop to facilitate an understanding of algorithms?
- Does your work put you in a unique position to understand algorithms differently? How so?
- How would you say your understanding of algorithmic systems is different from a dominant / practical understanding? What are your main criticisms of this dominant view?

(4) game designers

- Roughly, what are the some of the main stages of designing a videogame for development, beginning from scratch?
    - You can use a game you have developed or are developing as a case study.
- What kinds of conventions do you use? How do you balance between using conventions and breaking them?
- Is designing algorithms a task for game developers, or do game designers participate in algorithm design as well?
    - As a videogame designer, what understanding of programming, computation, or algorithms is sufficient?
- How would you say your understanding of algorithmic processes is different from that of a game developer?
    - From that of an algorithm developer?

- What is the appeal of low-resolution aesthetics?
- What kind of work goes into designing the game AI? How much was the AI playtested? What kinds of effects come out of it now?
- What does being able to break a game or exploit its rules bring to the gameplay experience?
- Why is a videogame a suitable medium for addressing the social impacts of gig labor and technocracy?

C) Interview Research Questions

1. How do participants in each setting consider the appearance and perception of algorithms in their work, if it all?
2. How do participants design for the appearance and perception of algorithms, if at all?
3. How do answers to these questions depend on the setting in question? What are common themes or practices across the settings?
4. Does the participants' work correspond to any of the paramediation patterns in Section 3? If not, does it evidence other patterns?
5. From the perspective of a theory of paramediation, how do the participants interact with paramediation in their work?
6. What is revealed and obscured by applying a theory of paramediation to understanding the participants' work?
7. To what extent is the participants' work generalizable to other settings of working with algorithms?

D) Gameplay Questions

Following 30 minutes of participant gameplay:

- How would you describe your experience playing the game?

- Did the elements in the game remind you of anything?
- Did the elements in the game remind you of any aspects of your own work?

After describing the theory of paramediation to the participant:

- Does the game inform your understanding of this theory at all?
    - What elements in the game are successful at doing so?
    - What elements in the game seem to contradict this theory?
- Does the theory of paramediation account for considerations in your own work, or not?

E) Theory Description Script

My goal in this study was to develop a theory about the social consequences of algorithms. I was specifically interested in how algorithms appear to human perception. Against the idea that algorithmic systems regulate behavior like a code of law, I proposed that they can govern behavior by arranging appearances, without restraining activity outright. I developed the game in an attempt to illustrate this.

F) Second Interview Questions

1. What problems or frustrations emerge during gameplay in each category?
2. According to participant actions and verbal reports, how does the gameplay in each category explain or demonstrate properties of paramediation to them?
3. How do the participants' specializations inform their gameplay or perceptions of the game?
4. How do participants relate their experience playing the game to their own work, if at all?

References

Adler, Nancy E, Thomas Boyce, Margaret A Chesney, Sheldon Cohen, Susan Folkman, Robert L Kahn, and S Leonard Syme. "Socioeconomic Status and Health: The Challenge of the Gradient." *American Psychologist* 49, no. 1 (1994): 15–24. https://psycnet.apa.org/doi/10.1037/0003-066X.49.1.15

Albarrán-Torres, César Alberto. "Encoding Chance: A Technocultural Analysis of Digital Gambling," 2014.

Alvarado, Oscar, and Annika Waern. "Towards Algorithmic Experience: Initial Efforts for Social Media Contexts." In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, 1–12. Montreal QC, Canada: ACM Press, 2018. https://doi.org/10.1145/3173574.3173860.

Amoore, Louise, and Volha Piotukh. "Life beyond Big Data: Governing with Little Analytics." *Economy and Society* 44, no. 3 (2015): 341–66. https://doi.org/10.1080/03085147.2015.1043793

Ananny, Mike. "The Curious Connection Between Apps for Gay Men and Sex Offenders." *The Atlantic*, April 14, 2011. https://www.theatlantic.com/technology/archive/2011/04/the-curious-connection-between-apps-for-gay-men-and-sex-offenders/237340/.

Ananny, Mike, and Kate Crawford. "Seeing without Knowing: Limitations of the Transparency Ideal and Its Application to Algorithmic Accountability." *New Media & Society* 20, no. 3 (2018): 973–89. https://doi.org/10.1177/1461444816676645

Andersen, Christian Ulrik, and Soren Bro Pold. *The Metainterface: The Art of Platforms, Cities, and Clouds*. MIT Press, 2018.

Anikina, Alex. “Algorithmic Superstructuring: Aesthetic Regime of Algorithmic Governance.” *Transformations*, no. 34 (2020).

———. “What Moves Non-Player Characters: Network Aesthetics in the Gamespace.” *Parallax* 26, no. 1 (January 2, 2020): 89–102. https://doi.org/10.1080/13534645.2019.1685785.

Antinori, Arije. “Terrorism and DeepFake: From Hybrid Warfare to Post-Truth Warfare in a Hybrid World,” In *Proceedings of the European Conference on the Impact of Artificial Intelligence and Robotics*, 23, 2019. https://doi.org/10.34190/ECIAIR.19.053

Avazpour, Iman, Teerat Pitakrat, Lars Grunske, and John Grundy. “Dimensions and Metrics for Evaluating Recommendation Systems.” In *Recommendation Systems in Software Engineering*, 245–73. Springer, 2014. https://doi.org/10.1109/TKDE.2005.99

Baeroe, K., and B. Bringedal. “Just Health: On the Conditions for Acceptable and Unacceptable Priority Settings with Respect to Patients’ Socioeconomic Status.” *Journal of Medical Ethics* 37, no. 9 (September 1, 2011): 526–29. https://doi.org/10.1136/jme.2010.042085.

Barlyn, Suzanne. “Strap on the Fitbit: John Hancock to Sell Only Interactive Life Insurance.” *Reuters*, September 19, 2018. https://www.reuters.com/article/us-manulife-financi-john-hancock-lifeins-idUSKCN1LZ1WL.

Bazemore, Andrew W., Erika K. Cottrell, Rachel Gold, Lauren S. Hughes, Robert L. Phillips, Heather Angier, Timothy E. Burdick, Mark A. Carrozza, and Jennifer E. DeVoe. “‘Community Vital Signs’: Incorporating Geocoded Social Determinants into Electronic Records to Promote Patient and Population Health.” *Journal of the American Medical Informatics Association* 23, no. 2 (March 1, 2016): 407–12. https://doi.org/10.1093/jamia/ocv088.

Beer, David. “The Social Power of Algorithms.” *Information, Communication & Society* 20, no. 1 (January 2, 2017): 1–13. https://doi.org/10.1080/1369118X.2016.1216147.

Beller, Jonathan. *The World Computer: Derivative Conditions of Racial Capitalism*. Duke University Press, 2021.

Berardi, Franco Bifo. “Precarious Rhapsody: Semiocapitalism and the Pathologies of the Post-Alpha Generation.” *London: Minor Composition*, 2009.

Bergen, Jan Peter, and Peter-Paul Verbeek. “To-Do Is to Be: Foucault, Levinas, and Technologically Mediated Subjectivation.” *Philosophy & Technology* 34, no. 2 (June 1, 2021): 325–48. https://doi.org/10.1007/s13347-019-00390-7.

Bogost, Ian. *Persuasive Games: The Expressive Power of Videogames*. MIT Press, 2007.

Bostock, Mike. “Visualizing Algorithms,” June 26, 2014. https://bost.ocks.org/mike/algorithms/.

Bowker, Geoffrey C, and Susan Leigh Star. *Sorting Things Out*. MIT Press, 1999.

Bratton, Benjamin. *The Revenge of the Real: Politics for a Post-Pandemic World*. Verso Books, 2022.

Bratton, Benjamin H. *The Stack: On Software and Sovereignty*. MIT press, 2016.

Bratton, William J. “Cops Count, Police Matter: Preventing Crime and Disorder in the 21st Century.” *The Heritage Foundation*, no. 1286 (March 27, 2018): 2018–03.

Bridle, James. “How Peppa Pig Became a Video Nightmare for Children.” *The Guardian*, June 17, 2018. https://www.theguardian.com/technology/2018/jun/17/peppa-pig-youtube-weird-algorithms-automated-content.

Brignull, Harry, Marc Miquel, Jeremy Rosenberg, and James Offer. “Dark Patterns - User Interfaces Designed to Trick People,” 2015. darkpatterns.org.

Bucher, Taina. "The Algorithmic Imaginary: Exploring the Ordinary Affects of Facebook Algorithms." *Information, Communication & Society* 20, no. 1 (January 2, 2017): 30–44. https://doi.org/10.1080/1369118X.2016.1154086.

Burrell, Jenna. "How the Machine 'Thinks': Understanding Opacity in Machine Learning Algorithms." *Big Data & Society* 3, no. 1 (January 5, 2016): 205395171562251. https://doi.org/10.1177/2053951715622512.

Campagna, Federico. *Technic and Magic: The Reconstruction of Reality*. Bloomsbury Publishing, 2018.

Chapman, Owen B., and Kim Sawchuk. "Research-Creation: Intervention, Analysis and 'Family Resemblances.'" *Canadian Journal of Communication* 37, no. 1 (April 13, 2012). https://doi.org/10.22230/cjc.2012v37n1a2489.

Chesney, Robert, and Danielle Citron. "Deepfakes and the New Disinformation War: The Coming Age of Post-Truth Geopolitics." *Foreign Aff*. 98 (2019): 147.

C.M. Hoonhout, Henriette. "Let the Game Tester Do the Talking: Think Aloud and Interviewing to Learn About the Game Experience." In *Game Usability: Advancing the Player Experience*. CRC press, 2008.

Coglianese, Cary, and David Lehr. "Transparency and Algorithmic Governance." *Administrative Law Review* 71, no. 1 (2018): 1–56.

Cohen, I. Glenn, Ruben Amarasingham, Anand Shah, Bin Xie, and Bernard Lo. "The Legal And Ethical Concerns That Arise From Using Complex Predictive Analytics In Health Care." *Health Affairs* 33, no. 7 (July 2014): 1139–47. https://doi.org/10.1377/hlthaff.2014.0048.

Cohen, Julie E. "The Surveillance-Innovation Complex: The Irony of the Participatory Turn." *The Participatory Condition (University of Minnesota Press, 2015, Forthcoming)*, 2014.

Cowgill, Bo, and Catherine Tucker. "Algorithmic Bias: A Counterfactual Perspective." *NSF Trustworthy Algorithms*, 2017.

Crawford, Chris. "Process Intensity." *The Journal of Computer Game Development* 1, no. 5 (1987).

Crawford, Matthew. "Algorithmic Governance and Political Legitimacy." *American Affairs* 3, no. 2 (2019): 73–94.

Crogan, Patrick, and Samuel Kinsley. "Paying Attention: Towards a Critique of the Attention Economy." *Culture Machine* 13 (2012).

Cuboniks, Laboria. *The Xenofeminist Manifesto: A Politics for Alienation*. Verso Books, 2018.

D'Agostino, Marcello, and Massimo Durante. "Introduction: The Governance of Algorithms." *Philosophy & Technology* 31, no. 4 (December 2018): 499–505. https://doi.org/10.1007/s13347-018-0337-z.

Daly, Kristen. "Cinema 3.0: The Interactive-Image." *Cinema Journal* 50, no. 1 (2010), 81–98.

Danaher, John. "The Threat of Algocracy: Reality, Resistance and Accommodation." *Philosophy & Technology* 29, no. 3 (September 2016): 245–68. https://doi.org/10.1007/s13347-015-0211-1.

Datta, Anupam, Shayak Sen, and Yair Zick. "Algorithmic Transparency via Quantitative Input Influence: Theory and Experiments with Learning Systems," *2016 IEEE Symposium on Security and Privacy*, (2016): 598–617. https://doi.org/10.1109/SP.2016.42

Dean, Jodi. "Communicative Capitalism: Circulation and the Foreclosure of Politics." *Cultural Politics* 1, no. 1 (2005): 51–74. https://doi.org/10.2752/174321905778054845

Debord, Guy. *Society of the Spectacle*. Bread and Circuses Publishing, 2012.

Deleuze, Gilles. *Cinema I: The Movement-Image*, U of Minnesota Press, 1986.

———. *Cinema II: The Time-Image*. U of Minnesota Press, 1986.

———. *Foucault*. U of Minnesota Press, 1988.

———. "Postscript on the Societies of Control." In *Surveillance, Crime and Social Control*, 35–39. Routledge, 2017.

Deleuze, Gilles, and Félix Guattari. *What Is Philosophy?* Columbia University Press, 1994.

Diakopoulos, Nicholas. "Accountability in Algorithmic Decision Making." *Communications of the ACM* 59, no. 2 (January 25, 2016): 56–62. https://doi.org/10.1145/2844110.

———. "Algorithmic Accountability: Journalistic Investigation of Computational Power Structures." *Digital Journalism* 3, no. 3 (May 4, 2015): 398–415. https://doi.org/10.1080/21670811.2014.976411.

———. "Algorithmic Accountability Reporting: On the Investigation of Black Boxes," 2014. https://doi.org/10.7916/D8ZK5TW2.

Diakopoulos, Nicholas, and Michael Koliska. "Algorithmic Transparency in the News Media." *Digital Journalism* 5, no. 7 (August 9, 2017): 809–28. https://doi.org/10.1080/21670811.2016.1208053.

Dourish, Paul. "User Experience as Legitimacy Trap." *Interactions* 26, no. 6 (2019): 46–49. dx.doi.org/10.1145/3358908

Drucker, Johanna. "Performative Materiality and Theoretical Approaches to Interface." *Digital Humanities Quarterly* 7, no. 1 (July 1, 2013).

Dwyer, Peter. "Conditional Citizens? Welfare Rights and Responsibilities in the Late 1990s." *Critical Social Policy* 18, no. 57 (November 1998): 493–517. https://doi.org/10.1177/026101839801805703.

Eslami, Motahhare, Aimee Rickman, Kristen Vaccaro, Amirhossein Aleyasen, Andy Vuong, Karrie Karahalios, Kevin Hamilton, and Christian Sandvig. "'I Always Assumed That I Wasn't Really That Close to [Her]': Reasoning about Invisible Algorithms in News Feeds." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 153–62. Seoul, Republic of Korea: ACM Press, 2015. https://doi.org/10.1145/2702123.2702556.

Eubanks, Virginia. *Automating Inequality: How High-Tech Tools Profile, Police, and Punish the Poor*. St. Martin's Press, 2018.

Federici, Silvia. *Caliban and the Witch*. Autonomedia, 2004.

Feenberg, Andrew. *Technosystem: The Social Life of Reason*. Cambridge, Massachusetts: Harvard University Press, 2017.

Feinberg, Melanie. "A Design Perspective on Data," 2952–63, 2017.

Floridi, Luciano. "Establishing the Rules for Building Trustworthy AI." *Nature Machine Intelligence* 1, no. 6 (2019): 261–62. https://doi.org/10.1038/s42256-019-0055-y

Flusser, Vilém. *Into the Universe of Technical Images*. U of Minnesota Press, 2011.

Foucault, Michel. *Security, Territory, Population: Lectures at the Collège de France, 1977-78*. Springer, 2007.

———. "Technologies of the Self," Vol. 18. Amherst, 1988.

Fricker, Miranda. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.

Friedman, Batya, and Helen Nissenbaum. "Bias in Computer Systems." *ACM Transactions on Information Systems* 14, no. 3 (1996): 330–347.

Fuchs, Christian. *Digital Labour and Karl Marx*. Routledge, 2014.

Fuller, Matthew, and Eyal Weizman. *Investigative Aesthetics: Conflicts and Commons in the Politics of Truth*. Verso Books, 2021.

Galloway, Alexander. "Theory Hot and Cold | Alexander R. Galloway," October 28, 2018. http://cultureandcommunication.org/galloway/theory-hot-and-cold.

Galloway, Alexander R. *Gaming: Essays on Algorithmic Culture*. U of Minnesota Press, 2006.

———. *Protocol: How Control Exists after Decentralization*. MIT press, 2004.

Garcia, Megan. "Racist in the Machine: The Disturbing Implications of Algorithmic Bias." *World Policy Journal* 33, no. 4 (2016): 111–17. https://doi.org/10.1215/07402775-3813015.

Gillespie, Tarleton. "The Relevance of Algorithms." In *Media Technologies*, edited by Tarleton Gillespie, Pablo J. Boczkowski, and Kirsten A. Foot, 167–94. The MIT Press, 2014. https://doi.org/10.7551/mitpress/9780262525374.003.0009.

Gillespie, Tarleton, and Nick Seaver. "Critical Algorithm Studies: A Reading List." *Social Media Collective* (blog), 2016. https://socialmediacollective.org/reading-lists/critical-algorithm-studies/.

Green, Donald P, Shang E Ha, and John G Bullock. "Enough Already about 'Black Box' Experiments: Studying Mediation Is More Difficult than Most Scholars Suppose." *The Annals of the American Academy of Political and Social Science* 628, no. 1 (2010): 200–208. https://doi.org/10.1177/0002716209351526

Groh, Matthew, Aruna Sankaranarayanan, and Rosalind Picard. "Human Detection of Political Deepfakes across Transcripts, Audio, and Video." *ArXiv Preprint ArXiv:2202.12883*, 2022. https://doi.org/10.48550/arXiv.2202.12883

Guattari, Félix. *The Guattari Reader*. Edited by Gary Genosko. Blackwell Readers. Oxford, OX, UK; Cambridge, Massachusetts, USA: Blackwell Publishers, 1996.

Haggerty, Kevin D, and Richard V Ericson. “The Surveillant Assemblage.” *British Journal of Sociology* 51, no. 4 (2000), 605–622.

Hague, Douglas C. “Benefits, Pitfalls, and Potential Bias in Health Care AI.” *North Carolina Medical Journal* 80, no. 4 (July 2019): 219–23. https://doi.org/10.18043/ncm.80.4.219.

Hajian, Sara, Francesco Bonchi, and Carlos Castillo. “Algorithmic Bias: From Discrimination Discovery to Fairness-Aware Data Mining.” In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, 2125–26, 2016. https://doi.org/10.1145/2939672.2945386

Hall, Mirko M, and Jürgen Link. “From the "Power of the Norm" to "Flexible Normalism": Considerations after Foucault.” *Cultural Critique*, no. 57 (2004), 14–32.

Hamburg, Margaret A., and Francis S. Collins. “The Path to Personalized Medicine.” *New England Journal of Medicine* 363, no. 4 (July 22, 2010): 301–4. https://doi.org/10.1056/NEJMp1006304.

Hardt, Michael, and Antonio Negri. *Empire*. Harvard University Press, 2000.

Hasen, Richard L. “Deep Fakes, Bots, and Siloed Justices: American Election Law in a’Post-Truth’ World.” *Saint Louis University Law Journal* 64, no. 4 (2019): 535–568.

Hui, Yuk. *On the Existence of Digital Objects*. U of Minnesota Press, 2016.

Hurst, S A. “Just Care: Should Doctors Give Priority to Patients of Low Socioeconomic Status?” *Journal of Medical Ethics* 35, no. 1 (January 1, 2009): 7–11. https://doi.org/10.1136/jme.2008.024323.

Ihde, Don. "Introduction: Postphenomenological Research." *Human Studies* 31, no. 1 (March 1, 2008): 1–9. https://doi.org/10.1007/s10746-007-9077-2.

Introna, Lucas D. "Algorithms, Governance, and Governmentality: On Governing Academic Writing." *Science, Technology, & Human Values* 41, no. 1 (January 2016): 17–49. https://doi.org/10.1177/0162243915587360.

Irani, Lilly, Janet Vertesi, Paul Dourish, Kavita Philip, and Rebecca E. Grinter. "Postcolonial Computing: A Lens on Design and Development." In *Proceedings of the 28th International Conference on Human Factors in Computing Systems - CHI '10*, 1311. Atlanta, Georgia, USA: ACM Press, 2010. https://doi.org/10.1145/1753326.1753522.

Joque, Justin. *Revolutionary Mathematics: Artificial Intelligence, Statistics and the Logic of Capitalism*. Verso Books, 2022.

Just, Natascha, and Michael Latzer. "Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet." *Media, Culture & Society* 39, no. 2 (March 2017): 238–58. https://doi.org/10.1177/0163443716643157.

———. "Governance by Algorithms: Reality Construction by Algorithmic Selection on the Internet." *Media, Culture & Society* 39, no. 2 (March 2017): 238–58. https://doi.org/10.1177/0163443716643157.

Kemper, Jakko, and Daan Kolkman. "Transparent to Whom? No Algorithmic Accountability without a Critical Audience." *Information, Communication & Society* 22, no. 14 (December 6, 2019): 2081–96. https://doi.org/10.1080/1369118X.2018.1477967.

Kücklich, Julian. "Precarious Playbour: Modders and the Digital Games Industry." *Fibreculture* 5, no. 1 (2005): 1.

Lala, Sayeri, Anastasiya Belyaeva, Maha Shady, and Molei Liu. "Evaluation of Mode Collapse in Generative Adversarial Networks," In *IEEE High Performance Extreme Computing*, 2018.

Langlois, Ganaele. "The Technocultural Dimensions of Meaning: Towards a Mixed Semiotics of the World Wide Web." York University and Ryerson University, 2008.

Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory*. University of Oxford Press, 2007.

———. *Science in Action: How to Follow Scientists and Engineers through Society*. Cambridge, Massachusetts: Harvard University Press, 1987.

———. "Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern." *Critical Inquiry* 30, no. 2 (January 2004): 225–48. https://doi.org/10.1086/421123.

Lazzarato, M. *Signs and Machines: Capitalism and the Production of Subjectivity*. Translated by Joshua David Jordan. Semiotext(e) Foreign Agents Series. Los Angeles, CA: Semiotext(e), 2014.

Lazzarato, Maurizio. "Immaterial Labour." *Contemporary Marxist Theory*, 1996, 77.

Lee, Min Kyung. "Understanding Perception of Algorithmic Decisions: Fairness, Trust, and Emotion in Response to Algorithmic Management." *Big Data & Society* 5, no. 1 (June 2018): 205395171875668. https://doi.org/10.1177/2053951718756684.

Lee, Min Kyung, Daniel Kusbit, Anson Kahng, Ji Tae Kim, Xinran Yuan, Allissa Chan, Daniel See, et al. "WeBuildAI: Participatory Framework for Algorithmic Governance." *Proceedings of the ACM on Human-Computer Interaction* 3, no. CSCW (November 7, 2019): 1–35. https://doi.org/10.1145/3359283.

Lessig, Lawrence. "Code Is Law." *Harvard Magazine* 1, 2000.

Liberini, Federica, Antonio Russo, Ángel Cuevas, and Ruben Cuevas. "Politics in the Facebook Era-Evidence from the 2016 US Presidential Elections," *CESifo Working Paper Series* 8235 (2020). https://dx.doi.org/10.2139/ssrn.3584086

Lievrouw, Leah A. "New Media, Mediation, and Communication Study." *Information, Communication & Society* 12, no. 3 (2009): 303–25. https://doi.org/10.1080/13691180802660651

Lyotard, Jean-Francois. *The Postmodern Condition*. University of Minnesota Press, 1984.

Malik, Kenan. "As Surveillance Culture Grows, Can We Even Hope to Escape Its Reach? | Kenan Malik." *The Guardian*, May 19, 2019. https://www.theguardian.com/commentisfree/2019/may/19/as-surveillance-culture-grows-can-we-even-hope-to-escape-its-reach.

Manovich, Lev. *The Language of New Media*. MIT press, 2001.

Marcus, Gary. "Deep Learning Alone Isn't Getting Us To Human-Like AI," *Noema Mag*, August 11, 2022. https://www.noemamag.com/deep-learning-alone-isnt-getting-us-to-human-like-ai.

Marmot, Michael. "Social Determinants of Health Inequalities." *The Lancet* 365, no. 9464 (2005): 1099–1104. https://doi.org/10.1016/S0140-6736(05)71146-6

Marx, Karl. *Capital: Volume I*. Penguin UK, 2004.

———. *Grundrisse: Foundations of the Critique of Political Economy*. Penguin UK, 2005.

Matzner, Tobias. "Beyond Data as Representation: The Performativity of Big Data in Surveillance." *Surveillance & Society* 14, no. 2 (2016): 197–210. https://doi.org/10.24908/ss.v14i2.5831.

———. "Opening Black Boxes Is Not Enough–Data-Based Surveillance in Discipline and Punish and Today." *Foucault Studies* 23, 2017, 27–45. https://doi.org/10.22439/fs.v0i0.5340

McLuhan, Marshall. "The Playboy Interview." *Essential McLuhan*, 1995, 233–69.

Mehrotra, Rishabh, Ashton Anderson, Fernando Diaz, Amit Sharma, Hanna Wallach, and Emine Yilmaz. "Auditing Search Engines for Differential Satisfaction across Demographics." In *WWW '17 Companion: Proceedings of the 26th International Conference on World Wide Web Companion*, 626–33, 2017. https://doi.org/10.1145/3041021.3054197

Menkman, Rosa. "Glitch Studies Manifesto." In *Video Vortex Reader II: Moving Images beyond YouTube*, edited by Geert Lovink and Rachel Somers Miles, 2011, 336–47.

Mittelstadt, Brent. "Automation, Algorithms, and Politics| Auditing for Transparency in Content Personalization Systems." *International Journal of Communication* 10 (2016): 12.

Mittelstadt, Brent Daniel, Patrick Allo, Mariarosaria Taddeo, Sandra Wachter, and Luciano Floridi. "The Ethics of Algorithms: Mapping the Debate." *Big Data & Society* 3, no. 2 (2016): 2053951716679679. https://doi.org/10.1177/2053951716679679

Miyazaki, Shintaro. "Algorhythmics: Understanding Micro-Temporality in Computational Cultures." *Computational Culture*, no. 2 (September 28, 2012). http://computationalculture.net/algorhythmics-understanding-micro-temporality-in-computational-cultures/.

Mol, Annemarie. "Ontological Politics. A Word and Some Questions." *The Sociological Review* 47, no. 1 (1999): 74–89.

Moore, Phoebe, and Andrew Robinson. "The Quantified Self: What Counts in the Neoliberal Workplace." *New Media & Society* 18, no. 11 (December 2016): 2774–92. https://doi.org/10.1177/1461444815604328.

Mozur, Paul. "Inside China's Dystopian Dreams: A.I., Shame and Lots of Cameras." *The New York Times*, July 8, 2018. https://www.nytimes.com/2018/07/08/business/china-surveillance-technology.html.

Nagel, Emily van der. "'Networks That Work Too Well': Intervening in Algorithmic Connections." *Media International Australia* 168, no. 1 (August 2018): 81–92. https://doi.org/10.1177/1329878X18783002.

Neyland, Daniel. "Bearing Account-Able Witness to the Ethical Algorithmic System." *Science, Technology, & Human Values* 41, no. 1 (January 2016): 50–76. https://doi.org/10.1177/0162243915598056.

Nicholas, Gabriel. "Shadowbanning Is Big Tech's Big Problem." *The Atlantic*, April 28, 2022. https://www.theatlantic.com/technology/archive/2022/04/social-media-shadowbans-tik-tok-twitter/629702/.

Nielsen, Jakob. "Usability Inspection Methods," In *Conference Companion on Human Factors in Computing Systems*, 413–14, 1994.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. NYU Press, 2018.

Ottinger, Gwen. "Making Sense of Citizen Science: Stories as a Hermeneutic Resource." *Energy Research & Social Science* 31 (2017): 41–49. https://doi.org/10.1016/j.erss.2017.06.014

Paglen, Trevor. "Harun Farocki (1944–2014)." *Artforum* (blog), February 6, 2015. https://www.artforum.com/passages/trevor-paglen-on-harun-farocki-1944-2014-50135.

Parisi, Luciana. "Instrumental Reason, Algorithmic Capitalism, and the Incomputable." *Multitudes* 62, no. 1 (2016): 98–109. https://doi.org/10.3917/mult.062.0098

Pasquale, Frank. *The Black Box Society*. Harvard University Press, 2015.

Peters, John Durham. "The Marvelous Clouds." In *The Marvelous Clouds*. University of Chicago Press, 2015.

Phelan, Chanda, Cliff Lampe, and Paul Resnick. "It's Creepy, But It Doesn't Bother Me." In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 5240–51. San Jose California USA: ACM, 2016. https://doi.org/10.1145/2858036.2858381.

Pinch, Trevor J, and Wiebe E Bijker. "The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other." *Social Studies of Science* 14, no. 3 (1984): 399–441.

Polack, Peter. "AI Discourse in Policing Criticisms of Algorithms." *Evental Aesthetics* 8, no. 1 (2019). Online: https://eventalaesthetics.net/wp-content/uploads/2021/03/EAV8_2019_Entire_Issue.pdf#page=57

———. "Beyond Algorithmic Reformism: Forward Engineering the Designs of Algorithmic Systems." *Big Data & Society* 7, no. 1 (2020). https://doi.org/10.1177/2053951720913064

Pollach, Irene. "What's Wrong with Online Privacy Policies?" *Communications of the ACM* 50, no. 9 (2007): 103–8. https://doi.org/10.1145/1284621.1284627

Rader, Emilee, and Rebecca Gray. "Understanding User Beliefs About Algorithmic Curation in the Facebook News Feed." In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, 173–82. Seoul, Republic of Korea: ACM Press, 2015. https://doi.org/10.1145/2702123.2702174.

Raghupathi, Wullianallur, and Viju Raghupathi. "An Overview of Health Analytics." *Journal of Health & Medical Informatics* 4, no. 3 (2013). https://doi.org/10.4172/2157-7420.1000132.

Rahwan, Iyad. “Society-in-the-Loop: Programming the Algorithmic Social Contract.” *Ethics and Information Technology* 20, no. 1 (March 2018): 5–14. https://doi.org/10.1007/s10676-017-9430-8.

Rancière, Jacques. *Dissensus: On Politics and Aesthetics*. Translated by Steve Corcoran. London; New York: Continuum, 2010.

Ratto, Matt. “Critical Making: Conceptual and Material Studies in Technology and Social Life.” *The Information Society* 27, no. 4 (2011): 252–60. https://doi.org/10.1080/01972243.2011.583819

Reigeluth, Tyler. “Recommender Systems as Techniques of the Self?” *Le Foucaldien* 3, no. 1 (2017): 1–25. https://doi.org/10.16995/lefou.29

“Request 20-5033 - NextRequest - Modern FOIA & Public Records Request Software.” Accessed August 3, 2022. https://recordsrequest.lacity.org/requests/20-5033.

Ribeiro, Manoel Horta, Raphael Ottoni, Robert West, Virgílio AF Almeida, and Wagner Meira Jr. “Auditing Radicalization Pathways on Youtube.” In *FAT* '20: Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 131–41, 2020. https://doi.org/10.1145/3351095.3372879

Rossiter, Ned. “Paranoia Is Real: Algorithmic Governance and the Shadow of Control.” *Media Theory* 1, no. 1 (2017): 88–102.

Rossiter, Ned, and Soenke Zehle. “The Aesthetics of Algorithmic Experience.” In *The Routledge Companion to Art and Politics*, 214–21. Routledge, 2015.

Rouvroy, Antoinette. “The End(s) of Critique: Data Behaviourism versus Due Process.” In *Privacy, Due Process and the Computational Turn*, 157–82. Routledge, 2013.

Rouvroy, Antoinette, and Thomas Berns. "Algorithmic Governmentality and Prospects of Emancipation." Translated by Liz Libbrecht. *Réseaux* 177, no. 1 (2013): 163–96. https://doi.org/10.3917/res.177.0163

Saba, Mariella, Tamika Lewis, Tawana Petty, Seeta Peña Gangadharan, and Virginia Eubanks. "From Paranoia to Power: Our Data Bodies Project 2016 Report," 2016. https://www.odbproject.org/wp-content/uploads/2016/12/ODB-Community-Report-7-24.pdf.

Shani, Guy, and Asela Gunawardana. "Evaluating Recommendation Systems." In *Recommender Systems Handbook*, edited by Francesco Ricci, Lior Rokach, Bracha Shapira, and Paul B. Kantor, 257–97. Boston, MA: Springer US, 2011. https://doi.org/10.1007/978-0-387-85820-3_8.

Sharon, Tamar. "Self-Tracking for Health and the Quantified Self: Re-Articulating Autonomy, Solidarity, and Authenticity in an Age of Personalized Healthcare." *Philosophy & Technology* 30, no. 1 (March 2017): 93–121. https://doi.org/10.1007/s13347-016-0215-5.

Shklovski, Irina, Scott D. Mainwaring, Halla Hrund Skúladóttir, and Höskuldur Borgthorsson. "Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use." In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2347–56. Toronto Ontario Canada: ACM, 2014. https://doi.org/10.1145/2556288.2557421.

Shneiderman, Ben. "Opinion: The Dangers of Faulty, Biased, or Malicious Algorithms Requires Independent Oversight." *Proceedings of the National Academy of Sciences* 113, no. 48 (2016): 13538–40. https://doi.org/10.1073/pnas.1618211113

Smith, Brian Cantwell. *The Promise of Artificial Intelligence: Reckoning and Judgment*. MIT Press, 2019.

Spradley, James P. *The Ethnographic Interview*. Waveland Press, 2016.

Stappers, P, and E Giaccardi. "Research through Design." In *The Encyclopedia of Human-Computer Interaction*, 2nd ed., 2017. https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/research-through-design.

Stiegler, Bernard. *Symbolic Misery. Volume 1: The Hyperindustrial Epoch*. Polity, 2014.

Strapagiel, Lauren. "This Researcher's Observation Shows The Uncomfortable Bias Of TikTok's Algorithm." *BuzzFeed News*, February 26, 2020. https://www.buzzfeednews.com/article/laurenstrapagiel/tiktok-algorithim-racial-bias.

Sun, Guo-Dao, Ying-Cai Wu, Rong-Hua Liang, and Shi-Xia Liu. "A Survey of Visual Analytics Techniques and Applications: State-of-the-Art Research and Future Challenges." *Journal of Computer Science and Technology* 28, no. 5 (2013): 852–67. https://doi.org/10.1007/s11390-013-1383-8

Sunstein, Cass R., and Richard H. Thaler. "Libertarian Paternalism Is Not an Oxymoron." *The University of Chicago Law Review* 70, no. 4 (2003): 1159. https://doi.org/10.2307/1600573.

Swan, Melanie. "Emerging Patient-Driven Health Care Models: An Examination of Health Social Networks, Consumer Personalized Medicine and Quantified Self-Tracking." *International Journal of Environmental Research and Public Health* 6, no. 2 (February 5, 2009): 492–525. https://doi.org/10.3390/ijerph6020492.

Terranova, Tiziana. "Red stack attack! Algorithms, capital and the automation of the common." *Effimera* (blog), February 12, 2014. http://effimera.org/red-stack-attack-algorithms-capital-and-the-automation-of-the-common-di-tiziana-terranova/.

Vaccaro, Kristen, Dylan Huang, Motahhare Eslami, Christian Sandvig, Kevin Hamilton, and Karrie Karahalios. “The Illusion of Control: Placebo Effects of Control Settings.” In *Proceedings of the 2018 CHI Conference on Human Factors in Computing*, 1–13, 2018. https://doi.org/10.1145/3173574.3173590

Vasudevan, Krishnan, and Ngai Keung Chan. “Gamification and Work Games: Examining Consent and Resistance among Uber Drivers.” *New Media & Society* 24, no. 4 (2022): 866–86. https://doi.org/0.1177/14614448221079028

Wark, McKenzie. “Gam3r 7H30RY.” *Cultural Politics* 2, no. 2 (2006): 213–24.

Wenz, Karin. “THEORYCRAFTING: Knowledge Production and Surveillance.” *Information, Communication & Society* 16, no. 2 (March 2013): 178–93. https://doi.org/10.1080/1369118X.2012.738695.

Westecott, Emma. “Game Sketching: Exploring Approaches to Research-Creation for Games.” *Virtual Creativity* 10, no. 1 (April 1, 2020): 11–26. https://doi.org/10.1386/vcr_00014_1.

Wittgenstein, Ludwig. *Philosophical Investigations*, Translated by Hacker and Schulte. Wiley-Blackwell: 2009.

Woodcock, Jamie, and Mark R Johnson. “Gamification: What It Is, and How to Fight It.” *The Sociological Review* 66, no. 3 (2018): 542–58. https://doi.org/10.1177/0038026117728620

Yanez-Gomez, Rosa, Daniel Cascado-Caballero, and Jose-Luis Sevillano. “Academic Methods for Usability Evaluation of Serious Games: A Systematic Review.” *Multimedia Tools and Applications* 76, no. 4 (2017): 5755–84. https://doi.org/10.1007/s11042-016-3845-9

Yeung, Karen. “Algorithmic Regulation: A Critical Interrogation.” *Regulation & Governance* 12, no. 4 (2018): 505–23. https://doi.org/10.1111/rego.12158.

———. "'Hypernudge': Big Data as a Mode of Regulation by Design." *Information, Communication & Society* 20, no. 1 (2017): 118–36. https://doi.org/10.1080/1369118X.2016.1186713

Yin, Xi, Xiang Yu, Kihyuk Sohn, Xiaoming Liu, and Manmohan Chandraker. "Towards Large-Pose Face Frontalization in the Wild." In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 3990–99, 2017. http://arxiv.org/abs/1704.06244.

Zhang, Richard, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 586–95, 2018.

Zhao, Zhengwei. "Analysis on the 'Douyin (Tiktok) Mania' Phenomenon Based on Recommendation Algorithms." Edited by G. Kierans, H. Liu, and E.H.K. Ng. *E3S Web of Conferences* 235 (2021): 03029. https://doi.org/10.1051/e3sconf/202123503029.

Ziewitz, Malte. "Governing Algorithms: Myth, Mess, and Methods." *Science, Technology, & Human Values* 41, no. 1 (January 2016): 3–16. https://doi.org/10.1177/0162243915608948.

Zuboff, Shoshana. "Big Other: Surveillance Capitalism and the Prospects of an Information Civilization." *Journal of Information Technology* 30, no. 1 (March 2015): 75–89. https://doi.org/10.1057/jit.2015.5.