

Jacob Kao and Anne Quaranto

Submitted May 2nd, 2025 in partial fulfillment of the requirements of the degree of Bachelors of Arts in Philosophy at Yale College.

Abstract

This paper interrogates the philosophical implications of social media, its underlying algorithms, and its implicit motivations through the application of social theories on selfhood. Drawing upon the work of George Herbert Mead and María Lugones, I argue that sustained participation in digital communities generates a "persona"—a durable, community-specific personality rooted in the social environment and the adoption of a "generalized other". While social media offers the same mechanisms of self-presentation and gestural feedback as offline socialization, the intervention of profit-driven algorithms introduces a unique set of affordances that distort the self-making process.

Unlike offline environments, social media platforms utilize algorithms designed to maximize engagement by prioritizing "System 1" stimuli—emotional, impulsive, and often inflammatory content—which skews the individual's perception of the generalized other. I contend that this environment forges a more direct connection between the symbolic and the material, resulting in online personas that are more salient and explicit than their offline counterparts. This asymmetry in salience leads to a "bleeding" of algorithmic values into the transcendental self, providing an explanatory framework for contemporary trends in radicalization, polarization, and the degradation of institutional values such as scientific accuracy and journalistic fairness. Ultimately, as the self identifies increasingly with its algorithmically structured persona, the values of real-life institutions lose their resonance and influence.

Keywords: Persona, Generalized Other, Ontological Pluralism, Algorithmic Affordance, Asymmetrical Salience, Profit-Driven Engagement, Institutional Degradation

I. INTRODUCTION

On October 28th of 2021, Facebook, a social media company with over 3 billion active monthly users, changed its name. The company, now 'Meta,' had a new "north star": the Metaverse.¹ This entirely virtual universe was to be populated by lifelike avatars and accessed through augmented and virtual reality devices.

The move was one that signalled Meta's vision of a surreal future; the kind prophesied by movies like *The Matrix* and *Ready Player One*. The public's reaction, though mixed at times, reflects a collective gut reaction: people do not buy this vision of the future.²

A future lived online, one where socialization occurs via technology—this is a

¹ *Meta*.

² Cordero Jr., "What a Wonderful World?"

future people hope to avoid, but it is a present that they already have. Today, people already live part of their lives on social media—two hours and twenty two minutes a day, to be exact.³ This time is not just spent scrolling and viewing—it is time spent actively engaging. Participation in social media contains a necessary component of self-presentation, evaluation, and assimilation with online communities. A loosely defined identity inevitably forms in tandem with this digital social space.

Spurred on by the complex and vast offerings of social media—never-ending posts, shares, likes, and direct messages—socialization online has become an increasingly intense and ubiquitous pastime. Moreover, as social media companies seek to earn from their users, they peddle a litany of stimuli, like a slot machine to a gambler, to pull and maintain users' attention.⁴ Algorithms now track what you like, how you like it, and how best to serve it to you. Though not yet an embodied Metaverse, the act of socialization online is just as immersive, identity-forming, and potentially harmful.

In the following pages, I interrogate the philosophical repercussions and implications of social media, its algorithms, and its company's motivations. Through the application of various social theories, I expose the mechanisms that make social media akin to, and more extreme than, real life socializing in important, structure-altering ways.

In my view, participation in social media generates a version of oneself—a kind of durable personality I coin *persona*. I expand on this later. Constrained by strategic and

profit-serving algorithms, individuals are afforded both unprecedented self-expression and, simultaneously, direct and unambiguous feedback. As such, an individual's online *persona* is conditioned by the nature of their perceived audience and the affordances of the algorithm.

In this paper, I argue that this unique online environment forges a more direct connection between symbolic and material, and offers affordances that shape—and disfigure—online *personas*—in a way that does not happen offline. These *personas* are more salient and more direct. As such, the judgment of offline and online *personas* is asymmetrical; as individuals become increasingly entrenched in online subgroups, algorithmic control, and inflammatory content, the resulting *personas* disproportionately affect the self. In part, this view aims to explain trends in radicalization, polarization, and the degradation of institutional values.

A VIEW OF SELFHOOD: THE PERSONA

To understand social media's impact on selfhood, one must examine how it intervenes in the self-making process. That is, how social media mimics and transforms the way that we come to be. Self-hood is a complex, iterative process that is rooted in the social. George Herbert Mead, an American philosopher and sociologist, attempts to pin down the specifics of this process in his 1934 book *Mind, Self and Society*. He asserts that the self follows from the apprehension of oneself as an object. Importantly, it is not in becoming a subject to oneself—it is the recognition that one is not merely the viewer, but also the viewed.

This act of self-recognition occurs, for Mead, within the confines of “a social

³ Howarth, “Worldwide Daily Social Media Usage (New 2025 Data).”

⁴ Busby, “Social Media Copies Gambling Methods ‘to Create Psychological Cravings.’”

environment or context of experience and behavior.”⁵ The process of developing self-hood is an intrinsically social one. It takes not only a theory of mind, but that theory applied to a set of individuals within a social environment that informs the way the act of object-perception takes place. In adopting this “generalized other”—the accumulated values, dispositions, and ways of seeing of a given community as understood by the individual—the notion of the self is directly informed by the social. Necessary in this process of accommodating the generalized other is the iterative act of responding to feedback.

Feedback is a nod, head-shake, or eye-shift. It is a handshake or failure-to, a word choice, an awkward silence. Gestures rule social convention, and they also set guidelines for how one ought to behave. According to Mead, in virtue of “a certain response in another...we shift from what we started to do because of the reply the other makes.”⁶ It is an active process of constantly updating one’s own actions to fit the perceived expectations and preferences of one’s audience.

Though gesture is the physical manifestation of how individuals receive and react to conveyed meaning, language is the most symbolically rich medium through which this happens. It is through an in-depth understanding of the complex features of language that people come to recognize the views of others, particularly with respect to themselves. As H.P. Grice argues in his landmark lecture “Logic and Conversation,” basic guidelines like the Maxims of Quantity, Quality, Relation, and Manner dictate norms around everyday conversation.⁷ When these norms are broken—when too much

information is provided, for example, or irrelevant details are included—violation of the norm itself signals content to the listener. All this is to say that language is a complex meaning-communicating medium. It is the primary means through which individuals situate themselves into the views of others, and thus recognize themselves as the object of those views.

The conception of others within a given environment, for Mead, culminates into this notion of the “generalized other.”⁸ This other, the amalgamation of a given social group, “gives to the individual his unity of self.”⁹ The understanding that the self has of the characteristics of his community—the way communication, perception, and meaning-making take place—is consolidated so that this other can influence the individual’s view of self. The concept much resembles Bernard Williams’ “imagined other,”¹⁰ or P.F. Strawson’s conception of the “Moral Community.”¹¹ In either case, through communication with community members, one gains a sense of the relevant attitudes of that community that can be applied to one’s self.

That communication informs features as basic as Grice’s conversational guidelines, and as subtle as racial biases implicit in meaning-making within this community. Importantly, Mead makes this point with respect to a given community. It is in exposure to a social environment that is homogenous in certain communicative practices, values, or dispositions that the individual comes to accommodate such features into his own conception of self and other.

⁵ Mead et al., *Mind, Self, and Society*, 2015. Pg. 2

⁶ Mead et al., *Mind, Self, and Society*, 2015. Pg. 2

⁷ Grice, *Logic and Conversation*.

⁸ Mead et al., *Mind, Self, and Society*, 2015. Pg. 3

⁹ Mead et al. pg. 4

¹⁰ Williams, *Shame and Necessity*. Pg 82

¹¹ Strawson, “Freedom and Resentment.” Pg. 92

This is not merely a sketch of how an individual adopts the biases of his environment. It is a much more profound assertion that one's deep-rooted understanding of self is dependent upon the environment in which that self is situated. Mead says, "He is taking over the morale of that society and is becoming an essential member of it. He belongs to it in so far as he does allow the attitude of the other that he takes to control his own immediate expression."¹²

Though Mead's discussion thus far has been on that of the self—and thus can be construed as a more overarching identity claim—I aim to direct his logic of developed self-hood to a weaker, more variable aspect of identity that more resembles personality—for my purposes, *persona*. Mead makes space for such a claim in his argument. "A person is a personality because he belongs to a community, because he takes over the institutions of that community into his own conduct. He takes its language as a medium by which he gets his personality..."¹³

He correlates "person" with "personality" because of how the features of personhood can and are rooted in the kinds of dispositions, communicative practices, and shared social meanings associated with personality. His intuition is supported by empirical findings, too. In her 1998 book *The Nurture Assumption*, Judith Rich Harris finds, through an emphasis on twin studies, that it is peer-group, not parents, that dictates how personality develops. She notes that, in instances of code-switching, individuals follow different norms of communication in the presence of parents than they do when around

friends, or the workplace. This is stark in the case of language, but extends naturally to behavior as well: "Language is the most conspicuous marker of the bicultural child's context-dependent socialization, but the assumption made here is that other variations in learned behavior are also selectively associated with the two different contexts."¹⁴

Mead's intuitive connection between socialized community and development of personality is supported by real empirical findings. The notion of personality-switching, here, is not a novel one. Most important, however, is Mead's explanation of the mechanisms through which community conditions personality. He asserts that the objectification of self through the lens of a generalized other is the means through which personality becomes aligned with environment. In doing so, he offers an explanation for how perversions of the generalized other, and of the self as an object, can impact one's identity.

Having pared back the explicit reach of Mead's logics to a discussion of *persona*—personality, fully embodied when occupying a given environment—I now will bring in another social theorist, María Lugones, and sketch how her view of identity can also offer relevant logics to my interrogation of social media. In brief, Lugones sets out to establish how theories of oppression can be liberatory from the subjective stance of the oppressed. She does this by sketching a kind of ontological pluralism—a plural view of identity that argues for multiple real ways of being. This plurality stems from the various communities to which the individual has access and in which the individual experiences life. The basic outline of this argument tracks well with Mead's notion of the self. However,

¹² Mead et al., *Mind, Self, and Society*, 2015. Pg. 5

¹³ Mead et al. Pg. 6

¹⁴ Harris, *The Nurture Assumption*. Pg. 462

Lugones goes further in arguing that, in instances of true embodiment within distinct communities—“realities”—it is a distinct self that exists within each community.¹⁵

The main thrust of her argument is that practical syllogisms—conclusions of practical action given one’s circumstances—can be produced only when one truly embodies the version of themselves created in response to a structuring environment. For example, a woman who is perceived through the lens of patriarchy her entire life has only the practical syllogism of continued existence within the constraints of that environment.¹⁶ She can alleviate those constraints only when she fully embodies an environment—say, a lesbian relationship—wherein she is perceived, treated, and allowed practical syllogisms that are completely distinct. For Lugones, distinct environments that contain truly distinct structures construe selves that act, enact, and animate their bodies with thoughts, feelings, and emotions unique to that self.¹⁷

Lugones sticks to ontological pluralism to present the strongest version of her view; one experiences the full act of self-recognition and self-understanding within a given reality, and this alone permits her practical syllogisms unique to that self. However, I argue that the logic of Lugones’ appeal to multiplicitous identity can be understood outside the confines of this absolutist, ontologically plural notion of selfhood. She herself, on occasion, appeals to an “in-some-sense the same person”¹⁸ even as she rejects the notion of a transcendental self that

Turner offers.¹⁹ I follow Turner in allowing for the possibility of this transcendental self even as I apply the logic of variable identity to my notion of *persona*.

I argue that the embodiment of a given environment, as Lugones describes, entails the production of a distinct, durable and characteristic *persona* that can produce practical syllogisms in response to the environment. She writes that “structures construct or constitute persons not just in the sense of giving them a facade, but also in the sense of giving them emotions, beliefs, norms, desires, and intentions that are their own.”²⁰ The characteristics she implies are necessary for person-hood can and do apply to my notion of the *persona*. The *persona* is not a mask, but a real lived and complex adoption of beliefs, norms, desires, emotions, and intentions. A *persona* is not a facade but a true and genuine way of interacting with oneself and others. Applying Mead’s argument for the generation of self-hood, the *persona* contains a distinct generalized other, and thus a novel view of the self as an object informed by that unique other.

I bring up Lugones because her notion of plural identity is helpful in understanding how the development of differing *personas* can create conflict and generate action. She appeals to a notion of liminality—of occupying a space between selves where each practical syllogism is frustrated, and thus structures are exposed—to do much of the heavy lifting. The limen, however, can be just as easily understood as an in-some-sense transcendental self, and thus is not necessary for my view.

In fact, mine is a kind of weak version of Lugones’ that I believe to track better with

¹⁵ Lugones, “Structure/Antistructure and Agency Under Oppression.” Pg. 57

¹⁶ Lugones. Pg. 55

¹⁷ Lugones. Pg. 57

¹⁸ Lugones. Pg. 61

¹⁹ Turner, *Dramas, Fields, and Metaphors*.

²⁰ Lugones, “Structure/Antistructure and Agency Under Oppression.” Pg. 60

common conceptions of selfhood and personality. People recognize oppressive and liberatory structures offered in their given environment through the different actions afforded their personas within those environments. The recognition of constraint and freedom is something more effectively realized in full embodiment—and the *persona* embodies a structure, environment, and mindset functionally as well as Lugones' self.

This short detour into the specifics of selfhood may appear tangential. However, the particular notion of a *persona*—of a durable set of characteristics rooted in the social environment and generalized other—is vital to interrogating the means and potential outcomes of our current engagement with social media.

III. SOCIAL MEDIA, PERSONA GENERATOR

This brings me to a full discussion of the nature of social media, how and why algorithms act the way they do, and the various effects engagement with these platforms has on the self. Only through a thorough analysis into the features of social media can I demonstrate the true impact it has, and will continue to have, on notions of self-hood, apprehensions of structure, and meaning-making.

First, in my discussion of social media, I refer to any website or app that enables its users to create and publish content, interact with others' content, and receive feedback in the form of likes, upvotes, comments, etc. Platforms such as Instagram, Facebook, Twitter (or X), and TikTok may come to mind, but my discussion extends to platforms such as Reddit, Letterboxd, Goodreads, Pinterest, and Discord. Any form of online media that attempts to entertain and also engage users through active participation is included in the discussion.

I begin by arguing that social media platforms offer a kind of socialization that resembles real life. Though their theories predate social media, Mead and Lugones have views that allow participation in social media to meet the requisite conditions for creating self-hood. Social media contains certain features of real life socialization such that the *personas* developed online should be understood as relevantly similar to those developed offline.

Mead argues that the self is created when an individual is capable of viewing himself as an object; as others in his environment would. Necessary, therefore, is the capacity to present oneself in a viewable way, and receive feedback through language and gesture in virtue of that presentation. Social media offers just this. Posting to social media is a form of self-presentation; even if that self is anonymous or otherwise unidentifiable. Implicit in the act of posting is the concept that the content is worth sharing to a given audience—otherwise the act of presenting to that audience is irrational. An implicitly social act, posting opens the individual up to the audience they intend to reach. This audience can upvote, like, dislike, comment, share, re-post, etc—this is platform dependent, but some form of interaction is universal across all social media. These interactions may be non-linguistic at times—an emoji, a “thumbs up,” etc—but always they communicate some attitude of the viewer. These are like the gestures that Mead speaks of as attitude-signalers in real-life conversation. Interactions may also be linguistic, and thus their similarity to real-life language is obvious. Comments, direct-messages, and replies all constitute common avenues through which an audience member can communicate their response to an individual's post. These follow

their own set of conventions that, when violated, offer their own various implicatures that individuals versed in communication on these platforms become aware of. For example, if a friend fails to comment on a post you tag them in, this may signal some kind of disapproval. If an individual that has just followed you likes an old post, this could signal a kind of interest. Each of these gestures, linguistic or not, offer the same subtlety of meaning contained in real life, in person conversation. Online, however, more objective, metric-based gestures such as ‘like’ counts make approval or disapproval more explicit. All the same, online gestures communicate content which contains meaning that can be interpreted to construe attitudes, dispositions, and preferences of one’s audience—just as is done offline.

Frequently, that audience is simply the general public, though it may also be a subreddit of Taylor Swift-hating people, or a private signal group chat planning the next attack against the Houthis. Whatever the case, individuals are cognizant of their audience and tailor their posts to fit that audience. Stefania Milan speaks to this act of fitting self with others online. She notes that “...the cloud allows individuals to tailor not only their participation but also their sense of belonging to a group,” and that this “self-expression” is voted upon by “selecting, emphasizing, and sharing content.”²¹ The norms of a given online community supervene over the content that is shared. Guidelines and values become implicit in the way individuals interact with each other, much as is the case offline.

As people become increasingly embedded within online communities—public or private—their notion of the generalized other of

that community becomes increasingly salient. Also, the practice of viewing self as object is made easy on social media. The notion of a ‘profile’ allows users to set a profile picture, a bio, and other information they feel relevant to present to others. Because this act of self-presentation is entirely manufactured, it becomes necessary for the individual to understand himself as a profile like any other. He must view his profile through the lens of how others might—otherwise there is no purpose in creating a profile, or interacting with social media at all.

All of this is to say, social media offers the same mechanisms of self-presentation, gestural feedback, and creation of generalized other that real life does. Given that the relevant features of persona-creation are present in social media, it follows that social media has the capacity to generate a *persona* in the way that Mead and Lugones reference. Though not physical, actual real-life presence—or embodied avatar—was never a prerequisite to self-creation in the first place. Milan asserts, in the context of resisting oppression, that, “thanks to social media, co-presence in a physical space is no longer a precondition for embodiment...in other words, one does not need to actually camp to undergo Occupy nor to join camps in foreign countries to experience a sense of transnational commonality.”²² She rightly asserts that embodiment—be that offline, or in the Metaverse—is not a prerequisite to participation, especially when the kind of self-presentation and gestural feedback necessary for community engagement is fulfilled sufficiently on social media. An individual can participate in a movement and assimilate to a group from behind a screen.

²¹ Milan, “When Algorithms Shape Collective Action.”
Pg. 6

²² Milan. Pg. 7

It is clear, then, that the *persona* created through social media checks the requirements set by Mead and Lugones for the generation of self. I go on to argue that social media actually goes much further. For a moment, though, consider what this sub-conclusion entails. A *persona* that is developed online views itself as an object, but the act of viewing is informed by the way that individual has seen it done within a community on social media. The resulting apprehension of self that this *persona* performs is guided by how others treat, react to, and post content online. This is all relative to a given community. Thus, in applying the concept of plurality that Lugones offers, an individual may have a number of distinct online *personas* in proportion to the number of communities that person is durably engaged and assimilated into. Much as individuals have *personas* for different real-life communities—church and school, for example—so too might they have them for a group chat with close friends and a Facebook group of car enthusiasts. It follows from this picture of *personas* rooted in online communities that the features of those communities, and the features of social media writ large will have profound impacts on the way notions of self-hood are formed for online *personas*.

IV. SOCIAL MEDIA GOES FURTHER

I assert that social media offers an experience that goes beyond the kind of interactions possible in real life. The affordances of the platform, the constancy of the stimulus, and the facilitation of motivated algorithms mean that the *personas* that people generate in online communities are distorted, salient, and potentially more impactful than any offline *persona*.

In 1982, James J. Gibson coined the term “affordances” to describe the possibilities for action that an environment offers an organism based on its capabilities.²³ For example, low-hanging fruit affords a child with opposable thumbs the ability to pluck it. If the fruit is high up, the child’s opposable thumbs are useless to this ends. Affordances are often discussed in reference to technology. A mundane example is how a ladder might afford the ability to climb high-up safely, or a door handle the ability to easily open and lock a door. The phone afforded long-distance communication, the text afforded a more convenient version of this, and so on. In the realm of social media, the affordances created are vast, complex, and completely transform the way that the social is conceived. In fact, Etlinger argues that “affordances as a field of possibilities are considerably more complex in algorithmic life than in a Gibsonian environment—actor relation in which there is only one actor.”²⁴

Far beyond the environments that Gibson and Mead conceived of, social media platforms, and their algorithms in particular, transform the way social life is conducted, and thus necessarily transform the formation of the self—and *persona*. For a true thoroughgoing investigation into the specifics of how social media has transformed the generation of self, one must turn to the structuring force of social media: algorithms.

In the simplest terms, “social media algorithms are designed to “reduce complexity brought about by information and interaction overload.”²⁵ This affords users the ability to

²³ Gibson, *The Ecological Approach to Visual Perception*.

²⁴ Etlinger, “Algorithmic Affordances for Productive Resistance.” Pg. 3

²⁵ Etter and Albu, “Activists in the Dark.” Pg. 71

interact with more individually-tailored information, and in turn makes the user experience conducive to spending more time on a given platform. This is because the currency of the social media economy is attention. The more time an individual spends on social media, the more time that platform has to serve up ads.²⁶ Algorithms are therefore necessary for a given platform's functionality for two reasons. Algorithms must be facilitate engagement and increase time spent, and they must also collect information on a user so as to personalize the advertising, and in turn charge more from advertisers. The question of why algorithms are the way they are has a simple answer: “maximizing user engagement will lead to a higher number of impressions on ads, which in turn will lead to profit from hosting these ads.”²⁷ Despite the intentional design of algorithms for the sake of profit, social media companies have little to no ability to predict how such algorithms will affect a user, outside of increased engagement. According to Arvind Narayanan, “...there is little human knowledge or intuition about what would make for a good recommendation that goes into the design of their algorithms. The algorithms are largely limited to looking for patterns in behavioral data. They don’t have common sense.”²⁸ Whereas much thought goes into the profitable measure of engagement and its maximization, there is little understanding of potential outcomes and externalities. Knowing how to keep someone online is very distinct from understanding the downstream effects of that behavior. Only in the practice of using and

participating in platforms that utilize such algorithms do social and interpersonal outcomes become apparent.

This is part of the reason why a philosophical interrogation of this subject is so vital. As a relatively new technology, social media and its long-term, downstream effects are relatively unknown, aside from immediate impacts on mental health and attention span that are already being noticed in the short term.²⁹ The effects of socializing in this way on conceptions of self, world and other are the kinds of features of human existence that will bear on any such long-term effects.

ALGORITHMS

Implications of this investigation aside, algorithms themselves are the difference-makers between off and on-line socialization. They impact the means through which meaning is shared, and thus skew users’ conceptions of the generalized other. As Milan makes clear, “[social media] is not a neutral vehicle: platforms not only carry but also ‘transform, translate, distort, and modify’³⁰ content and relationships—by, for example, metrifying interactions with the goal of altering the very same reactivity³¹ they incite.”³² Algorithms aim at increasing engagement and holding attention, so their motive is not to promote clarity or directness, but to serve themselves. It is in their best interest to create positive feedback loops of reactivity; to “transform, translate, distort and modify” so as to incite reactions that evoke reactions that evoke reactions, etc. Virality is an

²⁶ Kim, “Social Media Algorithms: Why You See What You See Technology Explainers.”

²⁷ Kim. Pg. 149

²⁸ Narayanan, Arvind, “Understanding Social Media Recommendation Algorithms.” Pg. 28

²⁹ Alaparthi, “Technology and Digital Media’s Impact on Attention Span in Teenagers and Young Adults.”

³⁰ Latour, *Reassembling the Social*. Pg. 108

³¹ Gerlitz and Lury, “Social Media and Self-Evaluating Assemblages.”

³² Milan, “When Algorithms Shape Collective Action.” Pg. 3

excellent example of this. If an individual creates a post that is particularly apt for a given audience, the algorithm will continue to serve it to members of that audience—and potential members—so as to keep users engaged. This is aside from the value, veracity, or importance of the post itself.

Users present something they believe others will appreciate and, in virtue of some characteristic about that post—not necessarily its value to the audience—it is boosted in the algorithm and the user receives feedback that their post was apt. In this way, perceived attitudes of the generalized other, and actual attitudes of that other become misaligned in a way that does not happen in real life. People construe attitudinal feedback as directly related to the content of their post—without necessarily thinking of how the algorithm exacerbates this effect. This can both bolster and discourage behavior that may not have normally been influenced without the intervention of an algorithm.

A potential counterargument, here, is that the content that algorithms promote is content that is already relevant to users, and that algorithmic promotion does nothing more than put the post in front of users' faces. If the post receives positive feedback, then it is because users appreciated the content, not because the algorithm forced them to.

Whereas this rebuke holds some truth, it forgets that an algorithm's core purpose is not to serve up what is most meaningful, truthful, or appreciated—only that which sparks engagement. This engagement need not be explicit, either. It can include, in the case of apps like TikTok, a measure of time-spent looking at a post: “Even if a user is passive and does not engage with a post, the algorithm

records the duration of time a user keeps the post on the screen as opposed to simply scrolling through.”³³ People who spend more time looking at outrageous claims, explicit imagery, or the like will find that the algorithm continues to serve them this content—whether they like it or not. This can do severe and profound damage to the individual. Their sense of the world and their conception of the generalized other skews towards the extreme that they are presented with. As Narayanan goes on to explain, the algorithm favors posts that evoke implicit feedback. These are the kinds that people are disproportionately served.

“The problem with implicit feedback is that it relies on our unconscious, automatic, emotional reactions: ‘System 1,’ rather than our rational and deliberative mode of thought: ‘System 2.’...a TikTok user might swipe past a video by a medical expert reminding people to get a flu shot because she doesn't look like the stereotype of a medical expert, and dwell on an angry video that they enjoy in the moment but regret later. By default, implicit-feedback-based feeds cater to our basest impulses.”³⁴

When users are judged on their system 1, they are judged on their most emotional, “basest” impulses. Sexual attraction, disgust, unconscious bias. Each of these is processed within system 1. Accordingly, the kind of content that individuals are offered, when the algorithm is trained on system 1, are more likely to be posts that evoke some reaction of, for example, sexual arousal.

The notion of intent, deliberation, and rational consideration are absent in system 1, yet these are the features of social interaction most

³³ Kim, “Social Media Algorithms: Why You See What You See Technology Explainers.” Pg. 150

³⁴ Narayanan, Arvind, “Understanding Social Media Recommendation Algorithms.” Pg. 36

conducive to collaboration and meaning-seeking. While real-life is privy to system 1 judgements within social interactions, these interactions are not made more common in virtue of some intervening agent. In a typical morning walk, you may encounter one or two stimuli that evoke a system 1 reaction—perhaps a car swerves towards you momentarily, or an attractive neighbor says hello. Compare this to the countless system 1 stimuli constantly served on Instagram or TikTok. The frequency of such stimuli influence how a rational person constructs his notion of the world and the generalized other. Consider the availability heuristic, which says that individuals base their perceived likelihood of an event on the number of examples of that event they remember.³⁵ This is an example of how short-cuts, developed for practical function, drastically impact the way people conceive of reality with respect to social media. The generalized other, if rooted in basic psychological principles like the availability heuristic, is sure to reflect system 1 disproportionately. For a *persona* constructed online, that other may be more hateful, lustful, or impassioned as a result.

The algorithm does not merely serve up more inflammatory content in service of continued engagement. It also facilitates the grouping of like-minded individuals who may chat and interact with each other more frequently than might two complete strangers. Friend recommendations and the formation of subgroups—like a subreddit on reddit, or Facebook group on Facebook—are relevant examples of this. This desire to form homogenous communities may also arise independently of algorithms. Seeking

like-minded people is natural, and all the more accessible in an online format. Physical distance, introductions, even language no longer stand in the way of accessing a community of people who think and act in the way want. This, too, has profound implications for how *personas* are created within these communities. If the generalized other is trained on a group of people that are entirely like-minded, then this *persona's* values, dispositions, and self-understanding will skew.

Now, it is true that surrounding oneself with likeminded people is not a uniquely online phenomenon. Real life affords this ability as well. Importantly, however, it does so much less easily, and in a way that is often circumstantial rather than intentional. A white man born into a community of only other white people will have a *persona* in this community that reflects the values of that given generalized other. However, the ability for a white person, born into a diverse community, to opt *into* an all-white community is afforded uniquely through the use of social media. Again, the ease of this affordance is not something that Mead could have anticipated when he described the generation of the self in 1934. The generation of that offline self is predicated on a nuanced environment, filled with variety upon which supervening similarity similarity is noticed and internalized as a characteristic of the generalized other.

Social media affords one far more control over the kind of *persona* they develop insofar as they can choose the community in which that *persona* is trained. Additionally, the group effect, which describes how views polarize in the echo chamber of a group, functions in full force within communities in social media.³⁶ While joining an anti-vax

³⁵ Kahneman, Slovic, and Tversky, *Judgment under Uncertainty*.

³⁶ Janis and Mann, *Decision Making*.

Facebook group may begin with naive curiosity, exposure to individuals who share this uncommon view, and the increasingly radical claims that these people make to stand out, can misrepresent and radicalize one's understanding of the topic. The ease of access of these communities to any individual makes them dangerous—and, compared to real life, markedly different.

At this point, one might argue that the effects I describe here—of polarization, radicalization, and increased exposure to system 1 stimuli—are not, in fact, unique to social media. These are on the rise in media writ large, today, and the fact that social media has followed suit is not uniquely impactful.

Though it is true that radicalization, exposure to inflammatory content, and the like are on the rise across all forms of media, I argue that involvement in social media is notably different. This is because the notion of a *persona* is an intrinsically social, self-presentational one. According to Lugones, it is only in “enacting [and] animating their bodies, having thoughts, feeling the emotions, in ways that are different in one reality than in the other” that a distinct self with durable “desires, character, and personality traits” comes to be.³⁷ This goes further than the picture of indoctrination that critics of news media often cite. This is the actual embodiment of those values, the presentation of a *persona* that presupposes those values and then is consistently rewarded for it. This is far more deep-rooted, and thus potentially far more impactful. It is only in the active participation in a different reality that one truly begins to form

the practical syllogisms in correspondence with those structures.

The notion of indoctrination sets up a simple presenter / viewer dynamic. The word is used in assertions like ‘schools indoctrinate children,’ or ‘news media indoctrinates political sects.’ The concept of indoctrination cannot account for the bidirectionality implicit in participation in social media. One not only views content that informs his perception of the generalized other, he adds to that content in virtue of acting in a way he believes is most presentable.

THE UPSHOTS OF A SOCIAL MEDIA PERSONA

It is worth, therefore, considering the implications, along logics like that of Lugones', of the development of these online *personas*. If an online *personas* exist only online, then this view has few offline implications. This is not the case, however. *Personas*, online or not, influence the self in much the same way that Lugones depicts selves contributing to liminality. Through the recollection of the self that was formed in a given reality, the individual can consider practical syllogisms true to that reality, but not true to others. This recollection occurs when *personas* are made evident to the self, be that a *persona* developed in a small groupchat or a *persona* created from a local book club.

I believe that a self inevitably recognizes—or has pointed out to them—the differences in behavior between *personas*. This may take the form of a frustrated practical syllogism, or of a slang word used in the wrong community. In any case, the self is capable of recognizing the differences in characteristics between *personas*, even as it continues to

³⁷ Lugones, “Structure/Antistructure and Agency Under Oppression.” Pg. 57

embody a given *persona*. As such, individuals gain some insight into the differences in their environments responsible for structuring such *personas*. Much like Lugones' framework, then, a multiplicity of *personas* exposes and enables the critique of structures, oppressive or not, of a given environment.

VII. CONCLUSION

The *personas* that are generated in virtue of participation in social media are relevantly similar, but importantly different, from those formed offline. Relevantly similar in the sense that they, too, can be apprehended, but importantly different in the extreme nature of their structuring environment. *Personas* that develop online are, as I have argued, more likely to generalize others along system 1-stimulating terms because of the intervention of algorithms in online socialization. Further, the social affordances of social media—the ease of joining groups, the act of providing unambiguous feedback in the form of a like or dislike, the creation of a ‘profile’ subject to self-objectification—mean that *personas* developed here are more explicit, and thus potentially more salient.

The augmented salience of the online *persona* to the individual has explanatory power in my view. I argue that features of social media make such a salience inevitable, and the resulting asymmetry of online and offline *personas* explains today's trends in radicalization, polarization, and the degradation of institutional values like journalistic fairness or scientific accuracy.

Consider the phenomenon of young white males entering an alt-right pipeline online, and adopting white-supremacist views in real life. The online *persona* of such an individual

enjoys the treatment of someone who is rewarded for slurs and ignorance. A *persona* that takes itself, informed by an other generalized on other white-supremacists, to be truly superior, to be worthy of better treatment, of greater access to power. When these individuals compare this *persona* to the one that exists in real-life communities, they apprehend oppression. The oppression of a school that does not allow for hate-speech. The oppression of society that says all are created equal. Where such institutions before would instruct about the nature of equality, they now are subject to questioning, because their values feel less important.

It is the fact that this individual has engaged online as if he is a white supremacist that allows him to realize the appeal of a white supremacist reality. He is not merely indoctrinated. The relationship is bidirectional. He presents in accordance to values he believes his audience shares, and receives support in virtue of that alignment.

This is an extreme example, one of the misapprehension of oppressive forces. Nothing about my view says that this is necessarily how online *personas* must be applied. It is true that online *personas* may be an individual's first time being treated equally, or being afforded the act of self-expression. The apprehension of real-life oppression can follow from such an online *persona*.

Insofar as online communities and content mirror that which is offered offline, *personas* will not be relevantly different. However, as I make abundantly clear in my discussion of algorithms, the *personas* that are developed online do so under algorithms that sort, present, and construe information in a way that values profit over all else. These algorithms

are excellent at sparking engagement, holding attention, understanding their users, all without drawing explicit attention to themselves.

These intervening algorithms make social media more salient, and thus facilitate an asymmetry with real-life that has worrying implications. It's the asymmetry that allows the apprehension of equality as oppression in the first place. An asymmetry that makes online communities appear more alluring, their values more integral, and their structures more self-evident.

As the self identifies increasingly with its social media *persona*, the values of that *persona* exert increased power. The values of real life that come into conflict with those developed online begin to degrade, as do their related institutions. Consider values of fairness in journalism, of accuracy in science, of aesthetics in art. Each of these is undermined by the kinds of content—inflammatory, and profit-driven—internalized and made normal through online socialization. In other words, these values of the offline *persona* resonate less with the self in virtue of the online *persona*'s greater appeal. As the gap between real, institution-upholding values and profit-making interests continues to widen, so too will this asymmetry in perceived selfhood. The artificially enforced, algorithmically structured *persona* will bleed into an increasingly artificial and algorithmic self as the offline *persona* loses its pull, its influence, and its agency.

During his re-naming of Meta, CEO of the company Mark Zuckerberg said that, despite the apparent differences between the Metaverse and Facebook, the mission of the company remained the same. "It's still about bringing people together," he said.³⁸ This act of bringing

together is what the company does. In its essence, it offers a platform and an algorithm that is socially satisfying for users—enough to serve them ads without turning them away. No matter how appearances change, the artificial facilitation of social interaction remains Meta's focus. The Metaverse is like social media. It only promises more comfort for online *personas* to live, socialize, and pay attention.

³⁸ *Meta*.

WORKS CITED

- Alaparthy, Kavin. "Technology and Digital Media's Impact on Attention Span in Teenagers and Young Adults." SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 29, 2024. <https://doi.org/10.2139/ssrn.4872178>.
- Atton, Chris. *The Routledge Companion to Alternative and Community Media*. First published in paperback. Routledge Companions. Abingdon: Routledge, Taylor & Francis Group, 2019.
- Bourdieu, Pierre. *Outline of a Theory of Practice*. Translated by Richard Nice. 1st ed. Cambridge University Press, 1977. <https://doi.org/10.1017/CBO9780511812507>.
- Busby, Mattha. "Social Media Copies Gambling Methods 'to Create Psychological Cravings.'" *The Guardian*, May 8, 2018, sec. Technology. <https://www.theguardian.com/technology/2018/may/08/social-media-copies-gambling-methods-to-create-psychological-cravings>.
- Cordero Jr., Dalmacito A. "What a Wonderful World? The Downside of the Metaverse Phenomenon." *Journal of Public Health* 46, no. 1 (February 23, 2024): e206–7. <https://doi.org/10.1093/pubmed/fdad177>.
- Edwards, Mary L., and S. Orestis Palermos, eds. *Feminist Philosophy and Emerging Technologies*. Routledge Studies in Contemporary Philosophy. New York, NY: Routledge, Taylor & Francis Group, 2024.
- Etter, Michael, and Oana Brindusa Albu. "Activists in the Dark: Social Media Algorithms and Collective Action in Two Social Movement Organizations." *Organization* 28, no. 1 (January 2021): 68–91. <https://doi.org/10.1177/1350508420961532>.
- Ettlinger, Nancy. "Algorithmic Affordances for Productive Resistance." *Big Data & Society* 5, no. 1 (January 2018): 2053951718771399. <https://doi.org/10.1177/2053951718771399>.
- Fardouly, Jasmine, and Lenny R. Vartanian. "Social Media and Body Image Concerns: Current Research and Future Directions." *Current Opinion in Psychology* 9 (June 2016): 1–5. <https://doi.org/10.1016/j.copsyc.2015.09.005>.
- Gerlitz, Carolin, and Celia Lury. "Social Media and Self-Evaluating Assemblages: On Numbers, Orderings and Values." *Distinktion: Journal of Social Theory* 15, no. 2 (May 4, 2014): 174–88. <https://doi.org/10.1080/1600910X.2014.920267>.
- Gibson, James J. *The Ecological Approach to Visual Perception*. 17th pr. New York: Psychology Press, 2011.
- Grice, H. P. *Logic and Conversation*. Foundations of Cognitive Psychology: Core Readings. Cambridge, MA, US: MIT Press, 2002.
- Harris, Judith. *The Nurture Assumption: Why Children Turn Out the Way They Do*. 2nd ed. Riverside: Free Press, 2009.
- Howarth. "Worldwide Daily Social Media Usage (New 2025 Data)." Exploding Topics, May 4, 2022. <https://explodingtopics.com/blog/social-media-usage>.

- Hunt, Robert, and Fenwick McKelvey. "Algorithmic Regulation in Media and Cultural Policy: A Framework to Evaluate Barriers to Accountability." *Journal of Information Policy* 9 (December 1, 2019): 307–35. <https://doi.org/10.5325/jinfopoli.9.2019.0307>.
- Janis, Irving L., and Leon Mann. *Decision Making: A Psychological Analysis of Conflict, Choice, and Commitment*. New York: Free Press, 1977.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky, eds. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge ; New York: Cambridge University Press, 1982.
- Kim, Sang Ah. "Social Media Algorithms: Why You See What You See Technology Explainers." *Georgetown Law Technology Review* 2, no. 1 (2018 2017): 147–54.
- Latour, Bruno. *Reassembling the Social: An Introduction to Actor-Network-Theory*. 1. publ. in pbk. Clarendon Lectures in Management Studies. Oxford: Oxford Univ. Press, 2007.
- Lua, Alfred. "Understanding the Instagram Algorithm: 7 Key Factors and Why the Algorithm Is Great for Marketers." *Medium* (blog), April 25, 2017. <https://medium.com/@shirleybwhalenca/understanding-the-instagram-algorithm-7-key-factors-and-why-the-algorithm-is-great-for-marketers-1777a937bbbf>.
- Lugones, Maria C. "Structure/Antistructure and Agency Under Oppression." *The Journal of Philosophy* 87, no. 10 (1990): 500–507.
- Mead, George Herbert, Charles W. Morris, Daniel R. Huebner, and Hans Joas. *Mind, Self, and Society*. The definitive edition. Chicago ; London: University of Chicago Press, 2015.
- Meta: Mark Zuckerberg Announces Facebook's New Name*, 2021. https://www.youtube.com/watch?v=Sya_ET05N7E.
- Milan, Stefania. "When Algorithms Shape Collective Action: Social Media and the Dynamics of Cloud Protesting." *Social Media + Society* 1, no. 2 (July 1, 2015): 2056305115622481. <https://doi.org/10.1177/2056305115622481>.
- Nagaiah, Aani. "Addressing the Mental Health Impacts of the Alt-Right Pipeline in Global Youth: A Medical Anthropology Perspective," January 20, 2024. <https://doi.org/10.58445/rars.893>.
- Narayanan, Arvind. "Understanding Social Media Recommendation Algorithms." Accessed April 23, 2025. <https://doi.org/10.7916/KHDK-M460>.
- Peterson-Salahuddin, Chelsea, and Nicholas Diakopoulos. "Negotiated Autonomy: The Role of Social Media Algorithms in Editorial Decision Making." *Media and Communication* 8, no. 3 (July 10, 2020): 27–38. <https://doi.org/10.17645/mac.v8i3.3001>.
- Scheinbaum, Angeline Close, ed. *The Dark Side of Social Media: A Consumer Psychology Perspective*. New York London: Routledge, 2018.
- Seymour, Richard. "The Machine Always Wins: What Drives Our Addiction to Social Media." *The Guardian*, August 23, 2019, sec. Technology. <https://www.theguardian.com/technology/2019/aug/23/social-media-addiction-gambling>.
- Strawson, Peter. "Freedom and Resentment." *Proceedings of the British Academy* 48 (1963): 187–211.

Turner, Victor. *Dramas, Fields, and Metaphors: Symbolic Action in Human Society*. Symbol, Myth and Ritual. Ithaca: Cornell University Press, 2018.

<https://doi.org/10.7591/9781501732843>.

Williams, Bernard. *Shame and Necessity*. Sather Classical Lectures, v. 57. Berkeley, CA: University of California Press, 2008.