

'FILTERED OUT' BUT NOT FORGOTTEN:  
HOW BLACK USERS CO-PRODUCE ALGORITHMIC IDENTITY ON TIKTOK

FALLON PARFAITE

A DISSERTATION

in

Communication

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2024

Supervisor of Dissertation

Dr. John J. Jemmott III

Kenneth B. Clark Professor of Communication and Psychiatry

Graduate Group Chairperson

Dr. John J. Jemmott III

Kenneth B. Clark Professor of Communication and Psychiatry

Dissertation Committee

Dr. Sarah Jackson, Presidential Associate Professor of Communication

Dr. David Lydon-Staley, Assistant Professor of Communication

Dr. Julia Ticona, Assistant Professor of Communication and Sociology

'FILTERED OUT' BUT NOT FORGOTTEN: HOW BLACK USERS CO-PRODUCE ALGORITHMIC IDENTITY ON  
TIKTOK

COPYRIGHT

2024

Fallon Parfaite

## **ABSTRACT**

### **'FILTERED OUT' BUT NOT FORGOTTEN: HOW BLACK USERS CO-PRODUCE ALGORITHMIC IDENTITY ON TIKTOK**

Fallon Parfaite

John J. Jemmott III

Despite recent theories that algorithms 'filter out' users with marginalized identities, identity, Black users are understudied in algorithmic identity research, particularly in how they navigate personalized, algorithmically mediated environments like TikTok. I address this gap by exploring the complex interplay between racial identity, user engagement, and algorithmic bias on TikTok's For You Page (FYP). This is a mixed methods dissertation consisting of three papers, each examining the relationship between racial identity and the FYP algorithm from distinct perspectives. The first paper investigates Black users' folk theories about how the algorithm interprets their racial identity and influences their engagement with the platform. Using Critical Techno-Cultural Discourse Analysis (CTDA), it uncovers how users perceive and respond to algorithmic biases, providing foundational insights into Black users' navigation of algorithmically mediated environments. The second paper highlights how racial identity centrality shapes perceptions of algorithm responsiveness, showing that users who view their racial identity as central tend to interpret the FYP algorithm as both responsive and insensitive. The third paper introduces the concept of algorithmic dissonance, synthesizing findings from both CTDA and survey research to capture the conflicting emotions and beliefs Black users experience in relation to the FYP algorithm's racial bias. Together, these papers advance the study of algorithmic identity co-production and critical digital studies by demonstrating how Black users navigate, interpret, and resist

algorithmic power, illuminating the complex interplay between structural biases, user agency, and self-concept within the algorithmic space.

# TABLE OF CONTENTS

<b>ABSTRACT.....</b>	<b>III</b>
<b>LIST OF TABLES .....</b>	<b>VII</b>
<b>LIST OF FIGURES .....</b>	<b>VIII</b>
<b>CHAPTER 1: INTRODUCTION.....</b>	<b>1</b>
WORKS CITED.....	6
<b>CHAPTER 2: BREAKING THE BLACK BOX: A CRITICAL ANALYSIS OF BLACK USERS’ ENGAGEMENT WITH THE FYP ALGORITHM.....</b>	<b>7</b>
ABSTRACT.....	7
INTRODUCTION .....	8
LITERATURE REVIEW .....	9
METHODS.....	20
ANALYSIS .....	24
DISCUSSION .....	46
CONCLUSION.....	50
WORKS CITED.....	54
<b>CHAPTER 3: EXPLORING ASSOCIATIONS BETWEEN PERCEIVED ALGORITHM RESPONSIVENESS, ALGORITHMIC AWARENESS, AND MEASURES OF RACIAL IDENTITY.....</b>	<b>57</b>
ABSTRACT.....	57
INTRODUCTION .....	58
LITERATURE REVIEW .....	59
EXPLORATORY ANALYSIS.....	72
DISCUSSION .....	89

CONCLUSION.....	93
WORKS CITED.....	94
<b>CHAPTER 4 : REFRAMING ALGORITHMIC RESISTANCE: EXPLORING ALGORITHMIC DISSONANCE IN BLACK USER ENGAGEMENT WITH TIKTOK’S FYP THROUGH MIXED METHODS ANALYSIS.....</b>	<b>97</b>
ABSTRACT.....	97
INTRODUCTION.....	98
LITERATURE REVIEW: USER AGENCY V. SYSTEMIC BIASES.....	100
ANALYSIS: FINDING ALGORITHMIC DISSONANCE.....	113
DISCUSSION: REFRAMING ALGORITHMIC RESISTANCE.....	121
CONCLUSION: THE END OF EDGE CASES.....	124
WORKS CITED.....	128
<b>CHAPTER 5: CONCLUSION.....</b>	<b>131</b>
<b>APPENDICIES.....</b>	<b>135</b>
APPENDIX A: SURVEY INSTRUMENTS.....	135
APPENDIX B: CENTRALITY SCALE FACTOR LOADINGS.....	138
APPENDIX C: BOX PLOTS PER VARIABLE AND Z-SCORE OUTLIER ANALYSIS.....	139

## LIST OF TABLES

Table 1: Correlation Matrix.....	73
Table 2:VIF Values.....	73
Table 3:Total Awareness and Total Responsiveness Regression Results .....	78
Table 4: PAR and AMCA Total Regression Results.....	79
Table 5: PAI and AMCA Total Regression Results.....	79
Table 6: Total Responsiveness with AMCA Dimensions Regression Results .....	81
Table 7:PAR as Outcome with AMCA Dimensions Regression Results .....	82
Table 8:PAI as Outcome with AMCA Dimensions Regression Results.....	83
Table 9:Centrality x Salience Interaction Term Regression Results .....	87
Table 10: Centrality and PAR Interplay Mediation Results .....	88
Table 11: Centrality and PAI Interplay Mediation Results.....	88

## LIST OF FIGURES

Figure 1: Q-Q Plot of Residuals for model with all predictor variables (ind. AMCA dimensions) .....	74
Figure 2: histogram of exhaustive model residuals.....	74



## CHAPTER 1: INTRODUCTION

The field of digital identity scholarship has shifted significantly alongside the evolution of social media platforms. TikTok's 'For You Page' (FYP) algorithm fundamentally reshapes user engagement with online content by using algorithms to create personalized feeds that are readily visible to users, as opposed to other platforms where algorithms often operate behind the scenes. Unlike social networks that rely on follower relationships or subscriptions to suggest content, the FYP continuously surfaces videos from unfamiliar creators based on inferred user interests (Barta & Andalibi, 2021; Abidin, 2020). This endless scroll of short-form content dynamically aligns with users' preferences, reshaping their self-concepts and online behaviors, and fostering deeper interaction with the algorithm. This form of engagement introduces a complex interplay between user identity and the FYP algorithm, challenging digital identity scholarship to evolve. Consequently, balancing the substantial impact of the FYP algorithm on user identity while recognizing user agency and critical engagement is a central challenge in understanding personalized online interactions. Identity scholars across disciplines address this by critiquing and expanding earlier notions of algorithmic identity that previously emphasized the deterministic role of algorithms in categorizing users (Cheney-Lippold, 2011).

Karizat et al. (2021) apply a knowledge co-production framework to examine how users and the TikTok FYP algorithm collaboratively shape user identity. This lens emphasizes the dynamic process through which identity is constructed, where both user actions and algorithmic responses contribute to an ongoing negotiation of digital identity. Through this framework, Karizat et al. introduce the Identity Strainer Theory, which reflects users' beliefs that the FYP algorithm actively "filters out" marginalized social identities. This filtering process, as perceived by users, leads to algorithmic representational harm, where algorithmic bias reinforces stereotypes and misrepresents marginalized groups, and algorithmic annihilation, where underrepresented identities are effectively erased or rendered invisible. By

highlighting these effects, Identity Strainer Theory underscores the risks of algorithmic curation for marginalized groups, calling attention to how algorithms may amplify structural biases and contribute to the misrepresentation or exclusion of certain identities within digital spaces.

In this dissertation, I build upon Karizat et al.'s approach, extending the concept of identity co-production from a lens into a phenomenon worthy of individual study. User perceptions of the FYP algorithm responding to their identity connects studies from various disciplines to form an interdisciplinary body of algorithmic identity co-production studies. Drawing on recent studies from digital self-making, human-algorithm interaction, and folk theory research through the dissertation, I position algorithmic identity co-production as an interdisciplinary research topic that encompasses both the structural power of algorithms and the agency of marginalized users. This approach underscores that even within systems that may attempt to "filter out" certain identities, users are not merely subjected to representational harm or annihilation; instead, they actively work within these constraints to maintain and assert their identities.

Users with racialized identities navigate a complex space where externally imposed social categorizations intersect with personal understandings of self and worldview (Sellers et al., 1998). Positioned at the center of the structure versus agency debate, these users confront both internal and external racial dynamics that shape their identities, both online and offline. Black users experience this tension uniquely as they engage with the FYP algorithm, encountering structural power dynamics tied to race while curating a personalized experience. I explore how Black users continue to interact with the algorithm, even when they perceive it as biased or misrepresentative of their identities. Moving from the concept of algorithmic erasure to active engagement, I argue and demonstrate that, despite experiencing algorithmic exclusion, Black users are not passive recipients of algorithmic outputs and are not simply 'annihilated'. Instead, they actively engage with the FYP in ways that challenge, negotiate,

and sometimes resist the algorithm's biases. By focusing on this dynamic interaction, I highlight the nuanced ways in which Black users exercise agency within an algorithmically mediated space—engaging in algorithmic resistance, self-representation, and identity negotiation. By explicitly examining Black users' interactions with and perceptions of the FYP algorithm, I bridge the gap between individual agency and systemic algorithmic bias, offering a nuanced understanding of the relationship between automated, personalized technologies and user identity. This research is driven by the following central question:

**How do Black TikTok users perceive algorithmic bias on the FYP in relation to their identity, and how does this shape their behavior?**

To address this study's research question, I employ a convergent mixed-methods approach, integrating qualitative and quantitative analyses (Wittink et al., 2006). Previous studies on FYP algorithm perceptions have used varied methodologies. For example, Karizat et al. (2021) employed in-depth interviews to understand user folk theories of the FYP algorithm's response to identity. Meanwhile, in algorithm studies, measures of algorithmic responsiveness adapted from interpersonal interaction research aim to assess how users perceive algorithmic acknowledgment of their identity (Taylor & Choi, 2022). Although these studies come from distinct academic fields, they directly cite one another and collectively contribute to the interdisciplinary body of algorithmic identity studies. By combining qualitative and quantitative methods, this dissertation provides a more nuanced understanding of how users perceive the interplay between their identity and the FYP algorithm, bridging the gap between empirical studies of human-algorithm interaction and critical socio-cultural theories, such as Identity Strainer Theory.

Furthermore, it is my contention to demonstrate the benefits of breaking methodological barriers dividing qualitative and quantitative methods. While having a specialized research area is necessary, topical boundaries do not necessarily have to apply to methodological approaches. Future critical digital studies interested in user perceptions of algorithms and AI (Artificial Intelligence) can benefit from breaking down methodological boundaries to merge relevant research from across disciplines, as I do in this dissertation with identity co-production.

### **Positionality Statement**

The researcher's positionality shapes all interpretations, including how their identity intersects with those of the participants and the subject matter (Charmaz, 2006). As a Black woman, avid TikTok user, and former content creator, I am uniquely positioned to conduct this research. My deep, daily engagement with the platform has fostered a close and enjoyable relationship with my FYP algorithm. While focusing on the algorithmic co-production of racialized identities, I do not intend to frame the FYP algorithm as inherently racist, rather I highlight the perceptions of users who are uniquely positioned to offer perspectives that push discussions of algorithmic identity forward. TikTok's FYP algorithm can be an awesome user experience. It is easier for users to share ideas and entertainment from diverse perspectives that they might not encounter within a traditional, static social network. This personalized experience is rich and multifaceted, reflecting multiple aspects of user identity, with race being one critical facet. By focusing on this important dimension, we can gain a better understanding of digital culture's evolution in light of the racial dynamics embedded in algorithmic systems.

## Dissertation Structure

My mixed methods dissertation is parsed into three papers that contribute to algorithmic identity co-production research. The first, entitled “Breaking the Black Box: A Critical Analysis of Black Users' Engagement with Algorithmic Bias”, employs Critical Techno-Cultural Discourse Analysis (CTDA) created by Brock (2016) to examine Black users’ folk theories of the TikTok FYP algorithm, exploring how these theories reflect users’ perceptions of racial bias and shape their engagement strategies.

Next, I take a quantitative approach to capture nuances of Black users’ perceptions of the FYP algorithm. “Exploring Associations between Perceived Algorithm Responsiveness, Algorithmic Awareness, and Racial Identity Dimensions” investigates the role of racial identity centrality and salience in shaping perceptions of algorithmic responsiveness and insensitivity, emphasizing how these dimensions influence users' sense of validation or exclusion on the platform.

Finally, I triangulate the findings from the first two papers in “*Reframing Algorithmic Resistance: Exploring Algorithmic Dissonance in Black User Engagement with TikTok’s FYP through Mixed Methods Analysis.*” This theory paper introduces the concept of *algorithmic dissonance* to describe the conflicting emotions and behaviors Black users experience when navigating perceived biases in the FYP algorithm. Using theories of racial formation, I argue that the ongoing agency vs. structure debate in critical digital studies is better understood as a *friction*—a tension between user agency and structural algorithmic influence that shape’s identity co-production in complex ways. Together, these papers deepen our understanding of algorithmic identity co-production and underscore the importance of examining algorithmic interactions through an interdisciplinary, critical lens that highlights marginalized user experiences.

## WORKS CITED

- Abidin, C. (2021). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103.  
<https://doi.org/10.5334/csci.140>
- Barta, K., & Andalibi, N. (2021). Constructing Authenticity on TikTok: Social Norms and Social Support on the "Fun" Platform. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–29. <https://doi.org/10.1145/3479574>
- Brock, A. (2016). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide Through Qualitative Analysis*. Sage.
- Taylor, S. H., & Choi, M. (2022). An initial conceptualization of algorithm responsiveness: Comparing perceptions of algorithms across social media platforms. *Social Media + Society*, 8(4).  
<https://doi-org.proxy.library.upenn.edu/10.1177/20563051221144322>.
- Wittink, M. N., Barg, F. K., & Gallo, J. J. (2006). Unwritten rules of talking to doctors about depression: Integrating qualitative and quantitative methods. *Annals of Family Medicine*, 4(4), 302–309.  
<https://doi.org/10.1370/afm.558>

## CHAPTER 2: BREAKING THE BLACK BOX: A CRITICAL ANALYSIS OF BLACK USERS'

### ENGAGEMENT WITH THE FYP ALGORITHM

#### Abstract

This paper investigates how Black TikTok users perceive and engage with the platform's For You Page (FYP) algorithm, examining identity co-production in algorithmically driven environments. TikTok's FYP algorithm curates highly personalized content, reshaping user engagement and influencing self-concepts. Yet, little research addresses how racialized users experience and navigate perceived biases in this context. Using Critical Techno-cultural Discourse Analysis (CTDA), I identify three modes of engagement—tactical, experimental, and critical—that Black users employ to interact with the algorithm. Findings show that users see the algorithm as both a source of representation and a site of bias. Tactical engagement involves subtle strategies to influence algorithmic outcomes; experimental engagement entails testing the algorithm to understand its biases; and critical engagement reflects a passive, vigilant awareness of potential harm. These engagement modes reveal Black users as active participants in the co-production of their algorithmic identities. This study extends algorithmic identity research by centering marginalized perspectives and positioning identity co-production as an interdisciplinary area for future exploration to prevent exacerbating the tension between algorithmic bias and user agency in increasingly complex media environments.

## Introduction

The integration of algorithms into TikTok's For You Page (FYP) fundamentally redefines user engagement with online content. Unlike other platforms that prioritize content from an extended social network, TikTok's FYP generates a personalized home screen populated with content from unfamiliar creators, based on inferred user interests (Barta & Andalibi, 2021; Abidin, 2020). The FYP facilitates continuous content discovery, presenting users with an endless scroll of short-form videos tailored to their preferences, interests, and behaviors. By shaping users' self-concepts and online behaviors, this algorithmic curation fosters a dynamic process of identity construction, where users influence and are influenced by the content they engage with. This interplay between user identity and the FYP algorithm challenges existing digital identity scholarship, pushing it to evolve and account for these complex interactions.

A critical challenge in examining personalized interactions like the FYP is balancing the impact of the algorithm with recognition of user agency—the classic structure-versus-agency debate. Scholars have long critiqued the traditional notion of algorithmic identity, which focuses on automated categorization and targeted advertising without adequately considering the user's active role in shaping identity (Cheney-Lippold, 2011). As personalized feeds like the FYP gain prominence, recent research has shifted to focus on user perspectives, leveraging users' heightened awareness of the FYP algorithm to reinsert agency into the concept of algorithmic identity. These studies seek to develop new theories that explain how users actively construct their identities in collaboration with the algorithm.

To further explore this tension between user agency and algorithmic bias, it is essential to focus on the experiences of racialized users, who remain underrepresented in algorithmic identity research.



While theories like Identity Strainer Theory suggest that marginalized users are “filtered out” by the algorithm, many continue to engage with and co-produce their identities on the FYP. These users, particularly those with racialized identities, navigate a unique digital space where external social categorizations shape how they perceive themselves and the algorithmic biases they encounter (Sellers et al., 1998). Situated at the center of the structure-versus-agency debate, racialized users face the dual task of navigating both the algorithm’s influence and the broader societal constructs that define their offline identities. Despite this centrality, limited research has examined how racialized users perceive and respond to algorithmic bias on the FYP, or how perceived bias affects their identity co-production practices and behaviors. Additionally, the extent to which these users feel compelled to modify their engagement to align their “algorithmized self” with their self-concept remains underexplored.

By focusing specifically on Black users, this study aims to illuminate how racialized users navigate and interact with the FYP algorithm, examining the behaviors and strategies that enable them to construct identities within an algorithmically influenced environment. Through an exploration of how Black users perceive and interact with the FYP in relation to their racial identities, this research seeks to build a nuanced understanding of the increasingly intricate relationship between personalized algorithms and user identity. This study is guided by the following central research question: *what are Black TikTok users’ folk theories about the FYP algorithm interplaying with their racial identity?*

## **Literature review**

In this section, I carve out space for algorithmic identity co-production as a distinct area of study to avoid continuing the tension between user agency and algorithmic power. I then explain why CTDA is

an advantageous method for studying user discourse specifically about the algorithm on which the discourse is taking place.

### ***Algorithmic Identity Co-production***

In this section, I argue that algorithmic identity co-production should evolve from a narrowly defined analytical framework into a comprehensive concept that encompasses the expanding body of research on the intersection of algorithms and user identity. TikTok's algorithm-centered user experience offers a distinctive context to examine identity formation beyond the individual perspectives of user agency and algorithmic power. Lack of interdisciplinary rigor reveals the ongoing tension between systemic algorithmic power and individual user agency in research concerning the fyp algorithm. The algorithmized-self framework emphasizes users' reflexive engagement with algorithmic reflections of their identity, tilting toward agency, the Identity Strainer Theory emphasizes the structural biases inherent in the FYP algorithm that influence user behavior. Together, these frameworks underscore the need for a more nuanced understanding of identity co-production within algorithmically mediated spaces.

The Identity Strainer Theory, introduced by Karizat et al. (2021), uses a co-productionist lens to highlight how users construct folk theories about the FYP algorithm, particularly beliefs that the algorithm filters out or suppresses content from marginalized identities based on race, body size, ability status, LGBTQ identity, and other social markers. This theory sheds light on user perceptions of systemic bias embedded within the FYP, suggesting that users are not passive recipients of algorithmic influence but rather actively interpret and respond to algorithmic filtering through their own social and cultural understandings. By framing these perceptions as co-produced knowledge, the Identity Strainer Theory

acknowledges that algorithmic identity is shaped by both user engagement and the algorithm's underlying logic. This approach reveals how marginalized users attempt to understand and navigate biases that affect the visibility and representation of their identities on the platform.

However, while the Identity Strainer Theory successfully identifies the presence of perceived systemic bias, it overlooks the lived experiences of those it claims are 'filtered out' or marginalized within the algorithmic space. Although the theory posits that marginalized identities face unique challenges and constraints, it does not fully examine how these users interact with or feel about being excluded from prominent spaces on the FYP, nor does it consider the psychological and social impacts of this filtering on users' sense of identity and belonging on the platform. By focusing primarily on the collective beliefs about systemic bias, the Identity Strainer Theory risks generalizing the experiences of marginalized users, presenting them as an abstract group affected by algorithmic filtering rather than as individuals whose personal experiences with exclusion and representation differ widely.

Moreover, this limitation restricts the theory's ability to capture the nuanced ways in which marginalized users may still assert agency despite feeling excluded. Those 'filtered out' might engage in unique forms of resistance or alternative modes of self-expression, such as intentionally curating content to subvert the algorithm's filtering mechanisms or creating niche communities that foster visibility beyond the FYP's constraints. By not delving into these specific interactions, the Identity Strainer Theory leaves a gap in understanding how marginalized users negotiate identity and representation on a platform where their visibility is perceived as constrained.

Expanding identity co-production from a narrow analytical lens into a broader body of research unifies theories that illuminate the intersection of algorithmic bias, identity, and agency. To fully address

the experiences of users affected by algorithmic filtering, the Identity Strainer Theory would benefit from an expanded focus on the complex interplay between perceived bias and individual responses to exclusion. A more holistic approach examining these dimensions would allow the Identity Strainer Theory to capture the diverse ways marginalized users experience, interpret, and respond to biases embedded in the fyp algorithm. This expanded approach offers a more nuanced understanding of algorithmic bias, not only as a perceived phenomenon but as a lived reality that significantly shapes user agency, identity, and social connections.

Self-making scholarship offers another aspect to this holistic approach. On TikTok, identity formation primarily occurs through interactions with the FYP rather than direct engagement with other users. As the FYP algorithm curates the content users see and engage with, it has a significant impact on how they experience the platform and construct their self-representations. Bhandari and Bimo (2022) introduce the 'algorithmized-self,' introduced by to describe how personalized recommendation algorithms influence self-representation and self-making. The 'algorithmized self' builds on and refines the earlier 'networked self' framework, which examined how social media platforms facilitate self-presentation and identity management (Papacharissi, 2011). Unlike the 'networked self,' which relies on connections within social networks, the 'algorithmized self' emphasizes a self-making process uniquely mediated by algorithmic curation that mirrors and reinforces complex facets of users' identities. This perspective contributes to understanding how algorithmic processes dynamically intersect with personal identity, suggesting that the algorithm not only reflects user preferences but plays an active role in shaping the digital self. The authors argue that user interactions with algorithmically curated content reinforce certain identity aspects, the algorithm serves as a tool for self-making, co-constructing digital

identities through personalized content aligned with users' unique perspectives and preferences. This process of "dual engagement" —where users simultaneously shape and are shaped by the algorithm— results in a continuous cycle of identity co-production.

The concept of the algorithmic crystal, introduced by Lee et al. (2022), enriches the idea of dual engagement by illustrating the multifaceted ways algorithms shape and reflect user identity through a "refractive" process. Rather than simply mirroring users' self-concepts, the algorithmic crystal suggests that algorithms fragment and highlight various aspects of identity, contingent on both user interactions and algorithmic design. In the context of TikTok's FYP, this means that the algorithm amplifies certain facets of identity while allowing users to encounter and connect with others who share similar attributes, even if these connections are brief and selective. This refractive approach underscores the increased presence of the algorithm in shaping user experience while still emphasizing user engagement as a central force in the co-production process.

However, individually, both the algorithmic crystal and the algorithmized-self frameworks lack the critical perspective offered by theories like the Identity Strainer, which foregrounds systemic biases embedded in algorithmic systems. While the algorithmic crystal model captures the complex ways users perceive aspects of themselves and others through algorithmically curated content, it does not adequately address how algorithms may fail to effectively reflect marginalized identities. Similarly, the algorithmized-self offers insight into how personalization algorithms shape self-representation by mirroring user identities, emphasizing user engagement and agency in identity co-production. Yet, like the algorithmic crystal theory, it lacks a critical lens on the inherent biases that privilege some identities over others. Such a lens is essential for understanding how algorithms not only mirror user preferences

but also reinforce systemic inequalities, often marginalizing certain identities in the process. Together, these frameworks illuminate aspects of identity co-production while underscoring the need for a more critical approach that addresses the biases underlying algorithmic mediation. By integrating these concepts within a broader framework that includes critical approaches, we gain a more nuanced view of algorithmic identity co-production that balances user agency with an awareness of the structural biases shaping digital experiences.

In conclusion, I propose broadening the concept of co-production from a narrowly defined analytical framework to a comprehensive term that encompasses research at the intersection of algorithms and user identity. I argue that algorithmic identity emerges as a form of knowledge co-produced by the user's cultural and cognitive commitments in tandem with the algorithm's cultural and material influences. Beyond the specific theories introduced by Karizat et al. (2021), I employ algorithmic identity co-production as a term that reflects relevant interdisciplinary research, acknowledging the expanding importance of algorithms and identity across diverse fields. The persistent tension between systemic algorithmic bias and user agency remains central within this broader literature, underscoring the ways algorithmic processes shape user identities, often reinforcing existing biases while simultaneously interacting with individual agency. While individual studies vary in their approaches, examining these theories collectively as identity co-production research provides a more balanced view—highlighting the structural power of recommendation algorithms in shaping self-concept while also recognizing the role of user agency and perceptions. This collective perspective advances our understanding of digital identity in the algorithmic age, offering a more nuanced approach to the complexities of identity formation within algorithmically mediated environments.

## CTDA and Folk Theorization

Folk theories, a popular approach to understanding identity co-production with algorithms, involve studying user perceptions of how an algorithm functions. Since social media algorithms are opaque and hidden 'black box' systems (Pasquale 2015), users develop folk theories based on their experiences and perceptions to understand how these algorithms personalize and curate their experiences on the platform (DeVito et al., 2017). However, because social media algorithms constantly evolve and are highly contextual, expanding studies of individual folk theories within the context of a specific platform is challenging. Building on their prior work, DeVito (2021) addresses this limitation by proposing the framework of the *process of folk theorization*, which is the study of how users repeatedly engage in folk theory creation in response to their evolving media environment. Furthermore, the author argues that:

By deepening understanding in the specific context of self-presentation, a universal, constant [66] social process which is known to be impacted by algorithmic mechanisms [24], we can evaluate folk theorization's general utility as a tool that can be leveraged to start building and encouraging enhanced algorithmic understanding via an extensible literacy in this area. (p. 2)

I interpret DeVito's argument as a call for a theoretical shift away from studying *which* folk theories users create about algorithms (i.e., why did Black TikTok users organize a protest?) toward understanding *how* and *why* they create folk theories through their interactions (i.e., how do Black users perceive the algorithm when it is not responsive to their self-concepts?). This approach will, in turn,

encourage research that builds "extensible" theories instead of individual folk theories that quickly become outdated due to the rapidly evolving social media ecosystem.

Performing Critical Techno-Cultural Discourse Analysis (CTDA) enables the centering of Black users while maintaining the theoretical rigor required by folk theorization (as opposed to individual folk theories). CTDA, developed by André Brock (2016), is an analytical technique that combines critical cultural theory with the examination of digital phenomena and artifacts. It aims to understand technology's material and semiotic nature within users' offline social and cultural practices. This approach unveils power relations in technology use, emphasizing the connection between technology, culture, and technological practices. Drawing from his foundational work on Black Twitter and frustrations with the lack of representation in digital cultural theory, Brock (2016) defines CTDA as:

A multimodal analytic technique for investigating Internet and digital phenomena, artifacts, and culture. It integrates an analysis of the technological artifact and user discourse, framed by cultural theory, to unpack semiotic and material connections between form, function, belief, and meaning of information and communication technologies (ICTs). CTDA requires the incorporation of critical theory—critical race, feminism, queer theory, and so on—to incorporate the epistemological standpoint of underserved ICT users to avoid deficit-based models of underrepresented populations' technology use. (p. 1)

CTDA is a valuable approach for examining identity co-production processes with algorithms due to its emphasis on the semiotic relationship between users and technological artifacts. This approach aligns with literature on identity co-production, where users identify aspects of themselves in the personalized



content recommended by the FYP algorithm, leading them to seek alignment between algorithmic representations and their self-concepts (Bhandari and Bimo, 2022; Lee et al., 2022). I will utilize CTDA to investigate this connection between user identity and the FYP algorithm further.

It is essential to understand what discourse analysis of sociotechnical systems looks like without this critical lens to demonstrate the importance of CTDA in digital studies. For instance, Cotter (2019) employs thematic analysis to examine how aspiring and established digital influencers on Instagram interact with pre-existing discourse and the algorithm to gain visibility. The study's theoretical motivations include entrepreneurialism, neoliberalism, and the recognition that influencers often achieve success through pre-existing privilege (Banet-Weiser, 2012; Duffy, 2017). Situated at the intersection of critical algorithmic and media studies, the study samples conversations between influencers in Facebook groups focused on sharing success strategies, defining Cotter's concept of 'conscious interaction' with algorithms. This research shifts the discussion from how Instagram influencers 'game' the system to how they 'play the visibility game' through conscious interactions with the algorithm and other expert users. The visibility game concept acknowledges the mutual influence of users, algorithms, and platform owners in shaping platform culture. It highlights that winning and losing in this game may perpetuate hegemonic ideologies and values.

Cotter (2019) concludes that digital inequalities have been linked to factors like socioeconomic background (Hargittai and Hinnant, 2008; Van Deursen and Van Dijk, 2014) and that some users are not aware of algorithmic ranking on platforms (Eslami et al., 2015; Rader et al., 2018), resulting in some Instagram users being unsure of how to play or even being unaware of the existence of the visibility game. The author writes:

This suggests that the game reinforces offline hierarchies of social privilege, with “winners” being those with greater access to social, cultural, political, and economic resources. Suppose the population of influencers does, indeed, represent a privileged few. In that case, their influence on social media culture—and culture more generally as it bleeds into “real” life—may perpetuate existing hegemonic ideologies and values. Future studies might investigate the digital inequalities embedded in the game and the resulting configurations of sociality on social media.

(p.15)

Cotter argues that some influencers may be excluded from the visibility game due to a lack of algorithmic literacy and that the influence of the 'privileged few' on social media may have residual effects on offline life. In contrast, CTDA would center the identities of the users in the Facebook groups to investigate their relationship to what they conceptualize as an influencer and why they turn to Facebook groups and self-proclaimed third-party gurus for tips on how to gain visibility on Instagram. Performing CTDA in this context would place more importance on the identities of the content creators from the groups analyzed in this project. CTDA might suggest that Instagram algorithms are designed to promote content that aligns with hegemonic narratives, regardless of a user’s algorithmic literacy or skill level. Although Cotter mentions that future research should investigate inequities that may be embedded in the game, further inquiry using CTDA that draws on critical Black studies literature could dismantle the game metaphor altogether.

I argue that CTDA flips the logic from focusing on privileged social media influencers impacting “real world” culture to understanding the culture that enables platforms to position visibility as a game.

In turn, applying CTDA in this context would highlight why the type of users in the Facebook groups conceptualize visibility on Instagram as a game and call for further inquiry into how other user groups outside of this type of influencer perceive and conceptualize visibility. This framing is essential in emphasizing why CTDA is vital to the theoretical and methodological shift from developing individual folk theories to studying the process of folk theorization. While ‘playing the visibility game’ can be understood as a unique folk theory that may or may not be exclusive to influencers, folk theorization may focus more intently on the process by which these users perceive visibility on Instagram. The way aspiring influencers form groups to gain knowledge about the algorithm can serve as an example of this behavior. Recognizing that other users may exhibit similar behaviors in different cultural contexts allows ‘playing the visibility game’ to become one example of such behavior, rather than the final concept.

Algorithms cannot be separated from their context, which presents a challenge when creating theories that expand beyond the context of one platform. However, the critical cultural approach of CTDA allows for a study to concentrate on one platform while developing extensible theories that can apply beyond that specific platform. For example, expanding on Simone Browne’s digital Black feminist work, Peterson-Salahuddin (2022) employs CTDA to investigate the ways Black TikTok creators enact ‘digital dark sousveillance’ by leveraging algorithmic gossip and technological savviness to engage in anti-racist discussion and spread anti-racist messages on the platform. The author underscores that digital dark sousveillance can be applied more broadly to other digital spaces and is not limited to TikTok specifically:

Thus, digital dark sousveillance provides a map and framework for examining how users necessarily play on medium affordance and use the medium to interact with

surveillance systems to both allow the voices of BIPOC to be heard while evading technological capture. However, how these practices are engaged will differ on each platform depending on its medium affordances and specific algorithmic biases. Future research should further examine how acts of digital dark sousveillance are enacted on other algorithmically informed social media platforms to understand the larger nexus of practices that make up digital dark sousveillance. In addition, I encourage future research expanding on these findings to engage different, user-centric methodologies, such as semi-structured in-depth interviews and ethnography, to explain further the thought processes that inform these practices that may not be as readily visible within the posted content. (p. 16)

The framework of digital dark sousveillance can be applied across multiple platforms and speaks to a broader 'nexus' within the sociotechnical system of power and surveillance, which the author contextualized in TikTok. Similarly, to move toward a better understanding of identity co-production with social media algorithms, we must develop robust frameworks and theories that are applicable across platforms and rooted in critical cultural work—especially as platforms become increasingly complex and continue to evolve as 'black box' systems. Current theories, such as the "algorithmized self," offer robust theoretical frameworks that I will pressure test in the context of Black TikTok users.

## **Methods**

The dataset for CTDA included 138 TikTok videos scraped using a third-party API hosted on Apify.com. Apify features APIs for many social media platforms created and maintained by web

developer communities. TikTok has an official API for researchers, but it requires a lengthy application, has a long wait period for approval, and does not provide better results than any third-party alternative (cite xx). The first scrape attempt was based on hashtags relevant to the discourse of interest such as #BlackTikTok, #BlackFYP, and #Blackalgorithm. However, the API did not return videos that were remotely related to Black users discussing the algorithm. TikTok research shows that hashtags are becoming less relevant in algorithmically curated feeds (Bhandari and Bimo, 2022). My failed scrape is a testament to the decreasing relevance of hashtags on a video platform as opposed to a text-based platform like Facebook or X (previously Twitter).

To pivot, I removed the hashtags and set the API to scrape videos by the search key words “Black fyp algorithm” and “black fyp”. The sample then became videos populated in the search results.

It was unclear from the documentation how the API fetched and ranked videos. I reached out to the creators of the API to confirm its sorting function, and they confirmed that it uses TikTok’s default search filter without any modifications. Although it is impossible to know exactly what blend of popularity and relevance metrics TikTok search uses to rank videos, the API is returning output from the default search parameters with no additional filters.

TikTok limits the number of requests that can be made from a single IP address within a certain time frame, so it is common for the APIs on Apify to time out despite the set number of videos in the request. I ran 8 requests between April 9-12, 2024. Some requests returned data for up to 60 videos while others returned 20 or less. Relevant metadata included in the scrape for each video were video descriptions (captions created by user), video URL, number of plays, number of likes, number of

comments, number of shares, country of creation, and timestamp. To ensure that the videos were relevant to the research topic, I visited each URL to watch the content. I defined relevant videos as videos featuring a Black-presenting, self-identified creator and mentioning the FYP algorithm. Videos not about the FYP algorithm and videos featuring white-presenting creators were not included in the final sample.

After cleaning the data, the final sample included 138 videos. The API could not scrape videos posted within a certain time frame, and TikTok's default search ranking evidently does not prioritize the most recent videos. Consequently, videos in the final sample span a wide range of time. The oldest video was posted in December 2020, but most were posted in 2022 through April 2024. The average number of likes for videos in the sample was 23,292 and the average play count was 102,183. The sample had a mix of popular and relevant videos. Some videos were only played 300 times while others were played over 100,000 times. Similarly, the average number of comments per video was 441, with some videos at 0 comments and others with thousands. Once all the data was clean to create the final sample, I combined the relevant videos from each scrape into one master spreadsheet for analysis.

### ***CTDA Approach***

CTDA is an advantageous method for understanding user perceptions of the FYP in relation to their racial identity because it allows the technology to be analyzed not only as a tool but as a cultural artifact. This dual perspective enables a deeper exploration of the power dynamics inherent in social media platforms, which in turn informs my interpretation of the discourse. My approach to CTDA

involved a two-part reflexive process that was both systematic and rigorous, allowing for a thorough examination of textual, visual, and audio content to uncover underlying themes, ideologies, and cultural narratives.

To begin, I performed inductive coding to systematically interpret the data, ensuring reliability and consistency in my analysis. Rather than simply consolidating the codes to identify themes, I analyzed the patterns in the codes and the ideologies they revealed, allowing my understanding to evolve as I engaged with the data. This approach was guided by the principle that discourse analysis should consider not just the language used but also the broader context and the role of language in constructing social realities. As I continued working through the sample, I used my codebook as a guide to revisit and refine my initial interpretations. This process allowed me to remain open to new insights and patterns, ensuring that my larger thematic findings emerged organically from the relationship between my interpretation of the underlying sociocultural dynamics and the discourse itself. By allowing the themes to develop naturally, rather than forcing them to fit pre-existing categories, I maintained sensitivity to the nuances and complexities of the data.

In the first phase of analysis, I coded 40% of the sample (56 videos) using an inductive approach. I employed structural coding to label videos with terms related to my research questions, creating a codebook to track my observations systematically. Given the limitations of qualitative analysis tools like NVIVO for video content, I used an Excel workbook to serve as my codebook, where I also included user comments to capture the full scope of discourse around each video. This approach ensured that I accounted for the most relevant and impactful comments, stopping when saturation was reached. I manually transcribed video quotes when automated transcripts were unavailable, and I assigned each

code a color to visually cue emerging patterns. Through this process, I developed 22 codes that captured various aspects of tone, context, and content within the videos.

## **Analysis**

It was extremely evident throughout the content that users perceive the algorithm to be biased toward their racial identities, but this bias presented itself in positive and negative ways. Although bias typically has a negative connotation, in the context of a personalized feed that is designed to be biased toward user preferences bias can also be considered positive. 'Bias' was one of the most frequently used codes in my codebook, coded as a literal mention of the word in the content. As I reflected on my own use of the term and engaged with the rest of the discourse, it became clear that the concept of bias describes a complex and nuanced experience.

On one hand, users felt validated when the algorithm seemed to "know" them and reflected their identity accurately, leading to content that resonated deeply. On the other hand, there was palpable frustration that the algorithm showed racial bias that reinforced stereotypes or failed to capture the multifaceted nature of their identity, leading to a feeling of being misrepresented or pigeonholed. This dual perception of bias highlights the intricate relationship between user identity and algorithmic behavior, where the lines between personalization, popularity, and bias are often blurred, resulting in both empowerment and alienation. Black users help many different perspectives about the ways the algorithm interplays with their identities and strategies based off those perspectives. In analyzing Black user's discourse surrounding TikTok's algorithm, I identified three modes of user engagement that capture the complex relationship between Black users and FYP algorithm: experimental engagement, tactical engagement, and critical engagement.



**Experimental engagement** refers to users conducting tests based on their own folk theories to expose algorithmic biases. **Tactical engagement** involves specific and rigorous strategies to influence or manipulate the algorithm, such as sharing strategies to 'fix' the algorithm or 'get the Black girl algorithm', aligning their algorithmic identities with their self-concepts. **Critical engagement** Critical engagement occurs when users interact with the For You Page (FYP) with a heightened awareness of how the algorithm shapes their experience, but in a more passive mode. For Black users, this reflects a form of critical algorithmic literacy that enables them to question and analyze the algorithm's behavior, rather than actively challenging or attempting to change it. This passive yet reflective engagement helps them recognize that the racial bias they encounter is not simply a technical flaw of the algorithm but a reflection of broader systemic racism embedded in sociocultural structures. By maintaining this critical distance, they share insights and strategies to avoid internalizing harmful biases, fostering a more informed but less confrontational approach to navigating the algorithm.

These types of engagement are not mutually exclusive; they can be concurrent and influence one another. For example, experimental engagement with negative feedback often results in critical engagement. Tactical engagement with positive feedback can ease the pressure of critical engagement. Perceptions of bias learned through the hyper awareness of critical engagement can be motivation for tactical and experimental engagement. Together, these forms of engagement offer a nuanced understanding of how users navigate and negotiate their identities within an algorithmically driven environment, balancing awareness of bias with proactive engagement.

Framing my larger themes as forms of engagement was a natural choice, as this theme was prevalent across many user discussions. The term 'engagement' is not only commonly used throughout

user discourse, but also conceptually captures the effort and attention users put into their FYP algorithms. This also fits the goals of CTDA by situating my interpretation of the discourse in the materiality of the algorithmic environment in which it exists. CTDA enables me to capture processes folk theorization rather than report individual ephemeral folk theories. Furthermore, through framing my analysis as forms of engagement with the FYP algorithm, I summarize beliefs and behaviors exhibited by Black users using concepts that are applicable across various contexts and demographics.

### ***Tactical Engagement***

Tactical Engagement occurs when users employ sophisticated methods to influence the For You Page (FYP) algorithm's output. Users share strategies and tactics via the algorithm's basic affordances as well as more effortful features and settings, such as the hidden "Not Interested" button and blocking keywords in content preferences settings. Many users believe the algorithm is racially biased because it reflects racial logic, but they also perceive it as shapeable through their behavior. Users often experience the algorithm as a digital extension of the prejudices present in the offline world. Tactical engagement also highlights a sense of agency among users who believe that, despite these inherent biases, the algorithm can be influenced, "hacked," or "fixed."

Many users shared tips and strategies for how to make their algorithm reflect their identities. In a video stitched to another creator asking how other users were manipulating their algorithms to see more Black creators, another creator shares how she changed her algorithm in 10 days:

Girl, let me help you one. Do not engage, okay? If you don't want to see them on your page, do not engage. Because I follow over 2,500 people and they are all black. And somehow if I even like one thing, okay? Even if I like one thing of the Chuckies, all of a sudden, my whole entire For You page is loose leaf. How? When I follow over 2,000, you know what I'm saying...Also, when you're favoriting sounds, don't favorite the sound unless it's under a black creator because a lot of those sounds are stolen, and we already know they already get pushed in the algorithm double. Their sounds get fame over our sounds when we're the originators of the sound. So make sure that when you like the sound, favorite the sound, whatever. It's from one of us, right? Also, the Not Interested button is your friend. Even if you are interested, if it's the wrong tone, not your vibe, hit that button. (User, 2024)

The discourse surrounding TikTok's algorithm reveals a complex process that Black users undertake to ensure their FYP reflects their racial identity and cultural values. While these users demonstrate a sophisticated understanding of how to manipulate the algorithm through tactical engagement, this process is far from effortless and, crucially, it is not permanent. The need for constant vigilance and ongoing effort underscores the inherent limitations and biases embedded within the algorithm, which persistently favor certain content over others, making personalization for Black users particularly challenging. The user highlights the strategic and conscious decisions made to see content from Black creators on the FYP. Her approach reveals several nuanced tactics aimed at influencing algorithmic identity to better align with their self-concepts.

First, the user employs selective engagement by intentionally avoiding interactions with certain content, such as not liking or engaging with specific creators or trends. This tactic leverages TikTok's algorithm to curate the For You Page (FYP) in a way that prioritizes content from Black creators while minimizing exposure to irrelevant or undesirable content. This conscious decision-making process reflects an understanding of how selective interactions can shape their FYP algorithm.

Next, the follow strategy involves following over 2,500 accounts, all of which are Black creators. This deliberate choice further shapes their FYP to reflect and amplify Black voices and content, helping to maintain a feed rich in culturally relevant and supportive material. By carefully selecting who to follow, the user creates a community that reinforces positive representation and aligns with their cultural identity. However, this strategy accounted for only one metric the algorithm considers. As a result, as the user notes that minimal engagement with different content (referred to as "Chuckies") can disproportionately affect their feed despite their monotonous following. Additionally, the user engages in sound curation by only favoriting original sounds from Black creators. This action counters the algorithm's bias that often benefits non-Black creators who use sounds originated by Black artists. Sounds are extremely important on TikTok. 'Sounds' refers to both official clips of popular music and user generated sounds (such as video voice overs) that are re-purposed by other users. By deliberately supporting original creators, the user ensures they receive appropriate visibility and credit, challenging the algorithm's default tendencies that marginalize certain voices.

The strategic use of the "Not Interested" button further refines the user's viewing experience. Even when potentially interested in certain content, the user may choose to reject it if it does not align

with their values or vibe. However, this button is not on the surface of the FYP and requires a long press to bring up the options menu. It adds an extra step to fine-tune the algorithm's recommendations, maintaining an FYP that closely matches the user's preferences. Another comment reads: "Yup I always hit not interested, they pop up on mine when I stay off for more than a day. But all I see is black content." This user expresses the daily efforts required to maintain a personalized algorithm that curates content from Black creators. Through these detailed strategies, the user demonstrates a high level of awareness about how TikTok's features and algorithms can be manipulated to better serve personal preferences and cultural representation goals. This creative use of TikTok's sound feature and engagement options exemplifies a proactive approach to digital self-curation.

Additionally, one user commented: "It took about 3 or 4 days, I realized engagement changed things. My fyp ain't never been this lit now," highlights the initial success that can come from deliberate engagement. However, the fact that it takes days of concerted effort to achieve a more relevant For You Page (FYP) points to the significant labor involved in this process. This labor is compounded by the reality that the algorithm is constantly recalibrating, as noted by another user who admits, "lol, I try not to F up my algorithm too much. some days I scroll past every single white man. that helps. some days I only follow poc." Another user said it takes ten days to change the algorithm: "Also like, fav & follow back every time they Black and do it about 10 days. Changed my whole algo #algorithm #black #fyp." These comments reflect the ongoing maintenance required to keep the algorithm aligned with their preferences, suggesting that any lapse in engagement can quickly undo the personalization efforts.

Moreover, the suggestion to focus on the "FOLLOWING" tab, as one user mentions, "You have to go to the 'FOLLOWING' tab. Sounds like you're in 'For You'. The ppl ur following have real content. Try switching it," indicates a workaround rather than a solution. While Black users can and do manipulate TikTok's algorithm to better reflect their identities, this process is laborious and impermanent. The constant effort required to maintain a personalized feed that aligns with their cultural values highlights the algorithm's inherent bias and its failure to provide equitable representation without significant user intervention. This ongoing labor underscores the systemic issues within the platform, where personalization for marginalized groups is not automatically granted but must be continuously fought for through tactical and intentional engagement. By carefully curating a multitude of strategies and detailed tactics, users found that they can alter their experience on the FYP, making it more reflective of their identities and values. This theme underscores the duality of the algorithm as both a reflection of existing societal issues and a system that can be manipulated to better serve the needs of Black users.

### ***Experimental Engagement***

Users engage in experimental engagement as a proactive approach to testing and challenging the TikTok algorithm, uncovering its biases, and understanding the factors it considers when determining content visibility on the For You Page (FYP). This type of engagement reflects a deliberate effort to navigate the algorithm's complexities and, in some cases, to expose its limitations. The goals of these experiments vary widely—some aim to test specific features like AI skin-tone detection, while others attempt to manipulate or "hack" the algorithm to maximize visibility.

One creator conducted an experiment by altering the shade of her skin in her videos to measure the impact on video performance. Speaking directly to her audience, she shared:

I've been wanting to talk about this because I've noticed a trend with my darker-skinned mutuals, like my Black mutuals, and how often our content gets suppressed. I tested this theory last week, and when I lightened my skin in videos, they went viral. It's frustrating. The algorithm is literally against us... I'm not the standard of beauty, and dark skin isn't either, but it's not enough just to follow Black creators. You've got to like, comment, and share. It's been really hard... And it's sad because I'm here to create content and just be joyful, but how can I do that when this app is literally against me?

This experiment, involving the use of a filter to lighten her skin, demonstrates how creators utilize platform features to probe the algorithm's behavior. Many content creators have expressed similar concerns about their content being suppressed or "shadow banned"—a situation where their content is hidden from the FYP without explicit notice or explanation, potentially due to their skin tone. These concerns are rooted in the belief that TikTok employs AI technology to detect skin tone and systematically suppresses content from darker-skinned creators.

The culture of 'going viral' on TikTok is distinct from that on other platforms, driven by an algorithm that can be unpredictable and opaque. Creators experiment with various features, language, and topics, understanding that each piece of content is treated independently by the algorithm,

regardless of overall engagement or follower count. This unpredictability means that one video might perform exceptionally well while another receives minimal views, often without clear reasoning. When a video does 'go viral,' it may not reach mainstream media but is instead featured on the FYPs of users whose interests align with that content, making virality highly subjective.

In the context of experimental engagement, virality can be seen within a specific “realm” of TikTok, aligned with the creator's niche and personal goals. The creator mentioned above identifies her niche as ‘book-tok.’ Even without clear metrics or a standardized definition of 'viral,' the experience of altering skin tone and observing fluctuating views reinforces perceptions of algorithmic suppression rooted in racist beauty standards and colorism.

These experiments contribute to a shared cultural understanding of how the TikTok algorithm interacts with racial identity. In response to another video discussing the algorithm’s potential to detect skin tone and suppress content from dark-skinned creators, one user commented, "I thought it was just me, that I was going crazy. I even thought I was shadow banned at one point." This comment highlights the uncertainty and self-doubt that often accompany experiences of algorithmic bias. Through experimental engagement, users seek clarity and validation, challenging the algorithm's opacity and asserting their understanding of its impact on their visibility.

Another significant example of experimental engagement within the Black TikTok community is the 'Black girls breaking the algorithm experiment,' initiated by content creator Toxic Chef (henceforth Chef). This large-scale experiment highlights the proactive and collective efforts of Black women to



challenge TikTok's algorithmic biases and advocate for greater visibility and representation. However, it also underscores the labor-intensive nature of these efforts and the ongoing struggle to maintain such gains in a platform not inherently designed to support marginalized creators.

Chef's experiment emerged in response to the widespread perception that TikTok's algorithm systematically suppresses content from Black creators, particularly those who do not conform to conventional beauty standards. This belief is supported by numerous anecdotal reports from Black creators who feel their content is unfairly downranked or 'shadow banned,' making it invisible to users without notice. The experiment aimed to counter this suppression by encouraging many Black women to engage with each other's content in a structured and deliberate manner—liking, commenting, sharing, and using specific hashtags to boost visibility on the For You Page (FYP). In a follow-up video, Chef reflected on the experiment's initial success:

All right, y'all, I'm back home, and I had a chance to analyze the data, and I have to say, it worked for me. I mean, I've never had a post get as many views, likes, comments, or shares. And I've been doing this social media thing for, like, five years, so that's a big deal. So I wanna say thank you, because you girls really showed up and showed out. You give a black woman an assignment, we're gonna get it done. Period.

Chef's reflection highlights the collective strength of Black women in amplifying marginalized voices on a platform frequently criticized for its lack of equitable representation. The increased engagement

metrics—more views, likes, comments, and shares—demonstrate the tangible impact of these coordinated efforts in disrupting the algorithm's inherent biases. However, this victory is bittersweet, as it reveals the exhausting and continuous labor required to achieve even a baseline level of visibility and representation. By describing their participation and success as "breaking" the algorithm, Chef suggests that the platform's design is inherently stacked against them, requiring extraordinary effort to overcome structural barriers that should not exist in the first place.

However, as the experiment progressed, Chef's frustration with the algorithm grew, highlighting the limitations and challenges of sustaining such efforts over time. In a later video, Chef expressed her concerns:

All right, y'all, it was all fun and games when I first started this experiment. It was literally like, a passing thought. But now I'm invested, and let me tell you why. Basically, the results are indicating that if you get a lot of engagement, your content gets pushed. But it still seemed as though my content wasn't being pushed as much as it should have been based on the interactions, which baffled me. Right. So we kept trying. We kept running more experiments. The more insight that I got, the less views that I was receiving. So I need to know, is it me that TikTok is suppressing, or is it just in general?


This quote reflects a deepening frustration with the algorithm's unpredictable behavior. Despite initial success, Chef observed diminishing returns on her content, even as engagement remained high. This

situation highlights the precarious nature of algorithmic manipulation—while users can influence the algorithm to some extent, its inherent biases and opacity often result in inconsistent outcomes. Chef’s growing concern about whether TikTok is specifically targeting her content underscores the emotional and cognitive toll that such efforts can take on creators who are constantly battling for visibility.

Chef’s proposal to use search engine optimization (SEO) tactics to sustain the experiment’s success marks a shift from experimental engagement to a more strategic interaction with the algorithm:

Some of my research indicated that SEOs, which is search engine optimization, is, like, a big deal for TikTok. So, let’s come up with a word that everybody puts in their videos. Slides off the screen but comes up if you search that term. What term can we use, guys? So, once we come up with this consensus, we’ll all just use that search term. And I recommend it’s not Black Girl follow train. It’s got to be something different so that the folks don’t suspect it and suppress it.

This shift highlights the need for more covert strategies to avoid detection by the platform’s moderation systems, which could suppress the coordinated effort if it becomes too visible. The use of SEO tactics reflects a deeper understanding of TikTok’s underlying mechanics and a recognition that sustained success requires continuous innovation to outsmart a system that is not designed with marginalized users in mind. Furthermore, she encourages the experiment’s expansion by using TikTok’s search page and other tactics aimed at influencing the algorithm.

The broader impact of these experiments within the Black TikTok community is evident in the conversation they sparked about collective action and the potential for lasting change. There were hundreds of videos from users who do not identify as creators posting videos of themselves to contribute breaking their algorithms. One comment on Chef's video encapsulated this sentiment: "Black girls got the algorithm algorithming !!!!Yall know how we do ." The phrase "the algorithm algorithming" playfully acknowledges the complexity and opacity of the algorithm while simultaneously asserting a sense of control and agency over it. It appeared all throughout the discourse, signaling the experience of achieving a feed that adequately reflects users' self-concepts. However, Chef's ongoing frustration also highlights the limitations of such efforts. Even with significant community engagement, the algorithm's unpredictable and opaque nature often undermines these victories, reminding users that their control over the system is tenuous and temporary.

The 'Black girls breaking the algorithm experiment' represents a powerful example of how marginalized communities can engage with and challenge algorithmic systems that often fail to reflect their racial identities in meaningful ways. Through coordinated efforts and innovative strategies, Black TikTok users like Chef and her followers demonstrate that it is possible to influence these systems and create a more equitable digital environment. However, the experiment also highlights the ongoing challenges of sustaining such efforts and the need for continuous adaptation and innovation in the face of an ever-evolving algorithmic landscape. Chef's growing frustration underscores the laborious and often precarious nature of these efforts, reminding us of the critical need for ongoing examination and community-led intervention in shaping the future of algorithmic systems.

## ***Critical Engagement***

Critical engagement refers to users' belief that they must maintain a critical stance by questioning and analyzing the algorithm's behavior to avoid internalizing its bias. Rather than passively accepting the algorithm's output, users are encouraged to critically assess their content's performance, understanding that success on the platform requires more than just good content—it demands strategic thinking, resilience, and a refusal to be discouraged by the algorithm's flaws. This theme underscores the need for creators to remain proactive and adaptive, taking control of their digital presence despite the algorithm's inherent challenges.

Many users expressed that the algorithm's inherent bias manifests through the promotion of problematic trends that target Black people, particularly Black women. One user observed, "When engagement is not hitting ... the best way to get it ramped up is to drag Black women, specifically Black American women in some kind of derogatory way." Other users express a deep sense of alienation, feeling that the culture of the app is inherently against them. For example, one user created a video with the text overlay: "Me, black gay alt, showing up on the FYP," as she walks into a room saying "hey y'all," only for the video to quickly cut to her screaming, "leave me the f\*ck aloooneee" under the text overlay "TikTok HQ." The description reads: "This app hates everything that I am." Similar sentiments are echoed across the platform, with many creators and users feeling that the algorithm is designed to work against them, leading to feelings of frustration, discouragement, and even a sense that the platform harbors hatred toward them.

Interestingly, despite these feelings of alienation, creators continue to participate on TikTok, sharing their experiences and critiques on the very platform they feel is biased against them. This reflects the complex, symbiotic relationship between user identity and the FYP. Even as they express frustration with a platform that seems to perpetuate biases based on their racial identity and other aspects of who they are, these users remain engaged, underscoring the intricate dynamic between their algorithmic identities and the algorithm that shapes it.

Non-creator users maintain critical awareness of the content that their algorithm's curate. One way that these users experienced critical engagement is through the belief that the FYP algorithm promotes problematic trends that encourage self-criticism among Black users. A particular trend involves Black users sharing self-deprecating thoughts about other Black people, accompanied by the text overlay "As a Black person, I'm not afraid to admit...". This trend's prominence on the FYP has led some creators to critically examine its impact, recognizing how the algorithm may be amplifying content that fosters division within the Black community. In response to this trend, one creator posted a video urging others to be vigilant about their engagement with such content:

Hey, Black people, I want you to pay attention to your views and your engagement on any video that you make about this subject matter, which has literally turned into us ripping each other to shreds on this app. And the reason I'm inviting you to do this is because I'm starting to notice a trend. Whenever there's a trend that involves us being especially cruel to each other, going back and forth, talking about each other, clashing, it's very easy for people to go viral talking about that. I'm

talking about accounts that are in suppression city, accounts that are in the 200 view jail. If you see videos on this subject matter, just look at the creator's views on this subject versus everything else that they talk about, and you will see a stark difference. You'll also notice that this is being pushed more heavily to your FYP even when you engage in other trending topics. I just want y'all to notice and pay attention to that, cause whenever we are on here being negative towards each other, suddenly we are getting released from Suppression City. Hmm.

This video highlights the creator's awareness of how the algorithm may be incentivizing and amplifying negative interactions within the Black community, rewarding content that involves intra-community conflict with greater visibility. The creator's call for vigilance underscores the need for Black users to be critically aware of how their passive engagement with such content could unintentionally perpetuate harmful stereotypes and reinforce the very biases they seek to combat. By drawing attention to the discrepancy in engagement metrics between content that fosters negativity and other types of content, the creator emphasizes the algorithm's role in shaping what goes viral, often at the expense of more positive or constructive discourse.

The recognition of this harmful trend extends beyond individual creators, as evidenced by user comments that resonate with the idea of digital disruption: "Trying to disrupt our digital communities like they've done our IRL communities. Yup. Damn that is a good catch." This comment draws a parallel between the disruption of online Black communities and historical efforts to destabilize real-life Black communities, suggesting that the algorithm's amplification of divisive content is part of a broader

pattern of marginalization. The awareness of this pattern fuels a collective frustration among users, many of whom express disillusionment with the platform: "As a Black person, this is the reason why I don't care if this app gets banned. They have favorites. I also notice when something is derogatory towards Black ppl it's never flagged against community guidelines."

Another video responding to the trend further critiques the self-critical behaviors that exacerbate its negative impacts:

We can start by learning about WHY we are stereotyped in this way and look at the historical events that could have led to this type of behavior instead of attacking and belittling our own community right from the get go. I know we want and need to get better, but this is NOT the way to do it. Always preaching 'open-mindedness' but instantly attack these individuals for how they are right off the bat instead of trying to understand WHY they're like that.

This response calls for a deeper understanding of the root causes of the stereotypes and behaviors being discussed, rather than perpetuating them through immediate judgment and criticism. The video suggests that by engaging in such self-critical content without questioning its context, users may be playing into the algorithm's biases, which thrive on conflict and controversy. The creator's critique also points to a broader issue of how the platform's design can exacerbate harmful dynamics within marginalized communities, where content that tears down rather than uplifts tends to gain traction. Further comments highlight the frustration with the widespread nature of these trends and the broader



consequences they have: "Thank you for posting about this!!!! And then the comments are a bunch of non-blacks thinking this is the green light to join in. Like what is the benefit of this trend? Where's the 'fun'?" ;"I had to come back to this video cuz my FYP is FLOODED with this trend smh." These comments underscore the pervasive nature of the trend and the concern that it invites participation from outside the Black community in ways that are harmful rather than constructive. The repeated appearance of such content on users' feeds, reflects the persistent influence of the algorithm in promoting divisive and problematic content. Together, these responses reflect a critical consciousness among Black users regarding the ways in which the algorithm can manipulate community dynamics, pushing trends that may ultimately be damaging. The emphasis on vigilance and the need to question one's own engagement with such trends highlights an ongoing awareness of the algorithm's tendency to prioritize racially biased content, calling for a more thoughtful and intentional approach to participation on the platform.

Alongside this awareness is a call for personal responsibility, where users are encouraged to balance their understanding of these biases with proactive strategies to navigate them. One creator articulates the frustration felt by many marginalized creators who struggle for visibility on the platform:

People's just general lack of empathy and analytical skills crack me up so let me get into details on here. Black, other POC, hijabi, plus-sized, disabled, paper-wig and curly wig cosplayers often get suppressed on here because we don't fit the standard of the algorithm on here. There's been articles and investigations showing this disparity, especially with Black and plus-sized creators on here—being paid less, not

being viewed enough despite most of its creative avenues of TikTok performed and created by those in the Black Diaspora. We are literally just fighting for equity so that our videos perform the same as those that fit the beauty standard. I need people to stop 'all lives mattering' this issue. Saying 'I boost any cosplay if I like it' just proves your bias towards eurocentric beauty standards since again that's the people who are performing the best on here as the rest of us struggle. Work to curb your biases and generally help equal the playing field for all of us. Equity is not 'desperation,' it should be the norm.

This user highlights the systemic challenges faced by creators who don't fit the algorithm's preferred mold. The creator calls out the disparities in visibility and compensation, emphasizing that equity should be the baseline expectation, not an exception. The call to "curb your biases" suggests that viewers, too, have a role to play in challenging the algorithm's tendencies by actively curbing their own internalized biases based on the algorithm's output. In another example, a user makes the argument that it's not the algorithm itself that's the problem, but rather the existence of racism in society. Many users aligned with the opinion that the algorithm itself is not inherently racist, but the users are, which is reflected in the algorithm's performance. In response to one video taking on this perspective, a commenter pushes back against this view, stating: "This is false and has been researched. It is indeed racism on a racist app using a racist algorithm. Look into TikTok, go read about why it was created, and how differently China's 'TikTok' is. This app was made to purposefully divide people, pls. don't gaslight Black women." This exchange illustrates the tension between recognizing broader societal issues and holding the platform accountable for perpetuating those issues. The commenter insists that TikTok's algorithm is not just a

reflection of societal racism but is actively designed to exploit and exacerbate divisions, particularly along racial lines. This perspective underscores the importance of not dismissing the role of the algorithm in perpetuating inequality, urging users to remain critical of the platform's underlying intentions and its impact on marginalized communities.

In contrast, another user reflects on the resilience and mindset required to succeed on the platform despite these challenges. In a video, the creator records herself setting up her tripod, replacing the original audio with a trending laugh track sound. The text overlay reads: "When are y'all gonna grow up and stop blaming the algorithm?" This video reflects the belief that while the algorithm may be flawed, creators must still take responsibility for their engagement and content strategy. The top comment on this video supports this dual approach, stating: "I think 2 things can be true at the same time: mindset is important and we do have to work twice as hard." This sentiment acknowledges the reality of the algorithm's biases but emphasizes that personal effort and a strong mindset are also crucial factors in achieving success. It reflects a pragmatic approach—recognizing the platform's flaws while also encouraging creators to take control of their content and engagement strategies. Significantly, many Black users have noticed shifts in the algorithm that impact the type of content they see and the engagement they receive. One creator reflects on a recent, positive change in their TikTok experience:

Don't know what I've done different recently, but my TikTok algorithm has switched up on me in the best way possible. The way I'm seeing all these Black girls up and down my TL that live normal lives and they make regular money and they have

regular jobs and they're just posting their content cause they love to do it. I'm so here for it. I don't know if it's just me. Before the New Year, every time I open this app, I was seeing all the big, top influencers. Someone named Michael's, the Alex Earls. And that's fine. Like, I love their content... But it's something about seeing girls that look like you, that are similar to you, that are motivated and always posting their content. No matter if they get 50 likes, 1,000 likes, they're just doing their thing.

This user's observation underscores the perception that algorithms can sometimes shift in ways that better align with their identity and interests. However, this shift also highlights the unpredictability of algorithms, which can suddenly start favoring content that resonates more completely with the user's identity without a clear explanation for what triggered the change. Yet, this positive shift is not without challenges. The creator acknowledges the emotional toll of consuming unrelatable influencer content and the difficulties of staying motivated to post content despite limited visibility on the FYP:

Like, opening this app and seeing the big-time influencers all the time can be so exhausting and draining and you can feel really lonely and behind when you watch their content. But lately, I've been seeing the girls that are right there with me doing their normal, regular stuff. One thing I'm realizing lately is it's okay to create content and not be an influencer. I think the word influencer is literally so. But some of us girls truly just love romanticizing our life and making content, taking videos and editing and doing the normal stuff. Like, truly, it's a hobby. One Thing I've had to

work on is not letting the internet make me feel like what I'm doing and where I'm at in life and the content I'm creating, the stuff I'm putting out is not relatable or it's not valuable. Because it truly is. Most of us are all in the same boat. And there's plenty of us that are watching your content, and we find it relatable and we find it motivating and we find it inspiring, and we really do love it. It is so hard to find dark-skinned girls or brown-skinned girls that truly just vlog their life and post it on the internet. We don't pop up on the For You page. Like, you really have to search and dive. Recently they've been up and down my TL and I'm just loving it. Girl, it is okay to romanticize your life. It is okay to take videos and put the exposure low and make it luxurious.

This reflection captures the delicate balance between navigating the algorithm's biases and embracing the value of personal content creation. It also emphasizes critical engagement from the perspective of both the creator and the user. For creators, they continue to make content that brings them joy regardless of their limited visibility. For users, they cannot rely on the FYP algorithm to surface Black creators and must actively search for content from 'normal people'. Overall, this quote exemplifies why it is important to distinguish between influencers and creators on TikTok. Videos from smaller creators, with '50 likes or 1000 likes,' create authentic content that is often overshadowed by popular influencers. Some smaller creators may aspire to become a full-fledged influencer, but at this point in their journey they are simply normal, everyday users who enjoy making content. Influencer culture has a toxic connotation that is 'exhausting' and 'draining'. Creator culture is much more open, encouraging more

users to contribute. TikTok is specifically designed for easy contribution, as reflected by their extensive in-app recording, sound, and editing options.

In summary, critical engagement involves a clear-eyed awareness of the algorithm's biases, coupled with a commitment to personal responsibility and resilience for both creators and users to remain critical of the algorithm, and it is a responsibility that not everyone perceives to take on. Still, while the platform may be biased, these creators demonstrate that with a keen sense of awareness, strong mindset, and community support, it is possible to carve out a space that reflects their identity and values to enjoy the experience of consuming relatable content.

## **Discussion**

In sum, these forms of engagement—tactical, experimental, and critical—not only illustrate identity co-production but also establish it as an interdisciplinary research area that demands cross-field attention. These engagement types reveal a more profound intentionality and reflexivity than typical interactions, such as merely consuming or selecting content. Users continually adapt their strategies in response to the algorithm's behavior, navigating and negotiating power dynamics within algorithmically driven environments. While algorithmic identity literature seeks to re-center user agency, many existing theories, particularly those related to TikTok's FYP, ultimately emphasize the algorithm's influence over user identity.

For instance, the Identity Strainer Theory, using a co-productionist lens, frames interactions as “algorithmic resistance”—deliberate behaviors that aim to produce outcomes distinct from those the algorithm would typically generate (Karizat et al., 2021). Similarly, Algorithmic Crystal Theory

conceptualizes user engagement as “strategic refinement” or “polishing,” wherein users curate their interactions to shape the algorithm’s interpretation of their self-concepts (Lee et al., 2022). This “polishing” process serves as a digital mirror, reflecting an identity that users perceive as accurate and empowering. Although both theories acknowledge users as active participants in their digital self-representations, they primarily depict these actions as responses to the algorithm’s influence, positioning the algorithm at the core of identity co-production.

By contrast, my approach reframes these interactions explicitly as user engagement, centering the user and their agency within the identity co-production process. By synthesizing literature across disciplines, I have established a space for identity co-production as a distinct phenomenon that underscores how users assert ownership of their digital identities through complex, intentional engagement patterns. This framing emphasizes that users are not merely reacting to the algorithm; they are engaged in a nuanced form of co-production that challenges algorithmic determinism, even within the constraints of systemic biases. By focusing on user agency—particularly among marginalized groups—my framework provides a more user-centered perspective that appreciates the depth of users’ participatory roles in shaping their digital experiences and identities. This understanding suggests that Black users interact with TikTok’s FYP algorithm in ways that transcend strategic refinement or algorithmic resistance, revealing a sophisticated form of identity co-production unique to algorithmically mediated spaces.

The concept of the ‘algorithmized self’ further explores how sociocultural systems shape self-making, building on Hearn’s (2010) argument that self-identity and self-presentation are increasingly tied to capitalist growth, which continually seeks new profit sources (Bhandari & Bimo, 2022). With the

"datalogical turn," self-identity is now embedded within a form of capitalism that relies heavily on the collection, generation, and trading of consumer data (Gregory et al., 2015). This system commodifies users' "desires, emotions, and expressions" as resources displayed online, integrating self-expression practices into a broader economic framework (Hearn, 2017, p. 63). The 'algorithmized self' reflects this shift, suggesting that TikTok users experience what Bhandari and Bimo call 'dual engagement,' interacting with both their algorithmically mediated personas and the platform itself:

TikTok users occupy the precarious position of dually engaging with both an external and internal entity; they engage with versions of themselves as mediated through the algorithm. While other social media platforms facilitate interaction with other egos through various methods, TikTok emphasizes interaction with the personalized algorithm, which repeatedly confronts users with aspects of their personas (p. 9).

I interpret this 'dual engagement' as a form of identity co-production. Black TikTok users, in particular, demonstrate awareness of this dual engagement process, incorporating it into their self-making practices. Through sophisticated interactions with the platform, they become acutely aware of an 'external' entity—the algorithm—that influences their digital identity, leading them to perceive that the algorithm itself exhibits biases aligned with their racial identities. This awareness goes beyond simply recognizing that user inputs can influence the algorithm. Instead, these users engage in a dynamic process of identity co-production, informed by an understanding of how their identities are datafied and processed by the algorithm, often in ways that reinforce systemic biases.



Moreover, these forms of engagement—tactical, experimental, and critical—not only illustrate identity co-production but also serve as clear evidence that the "datological turn" is now fully realized. This turn represents a shift where self-identity and social engagement are no longer confined to offline or even strictly personal realms but are intrinsically woven into data systems that capitalize on user behavior and expression. Users are now acutely aware that their digital selves are being converted into data points, commodified, and traded within larger algorithmic and economic structures. This heightened awareness propels users to engage with algorithms in ways that go beyond mere interaction, positioning them as both subjects and co-creators within a system where identity itself is a form of currency.

In this context, the various forms of engagement users employ on platforms like TikTok reflect more than individual attempts at self-expression; they signal a deep-seated awareness of the algorithm's role in shaping self-identity, one that acknowledges the socio-economic forces embedded within these platforms. Tactical engagement, for example, shows that users recognize how specific behaviors can influence algorithmic outputs, enabling them to strategically present aspects of themselves that align with or subvert the algorithm's perceived preferences. Experimental engagement reveals that users not only understand but test the boundaries of these data-driven spaces, exploring how different interactions produce different versions of their digital self. Critical engagement, on the other hand, underscores users' awareness of the algorithm's biases and limitations, as they intentionally push back against or navigate around these biases to assert a more authentic self-representation.

Together, these engagement practices demonstrate that users are not merely passive recipients of algorithmically curated content; they actively participate in and shape the algorithmic processes that,

in turn, shape them. This feedback loop epitomizes the datalogical turn: identity formation is no longer simply a personal or social process but a co-produced, data-driven phenomenon embedded within and influenced by complex social, cultural, and technological infrastructures. Users' conscious manipulation and negotiation within these systems reveal a sophisticated understanding of how their data is used and the stakes of their digital self-presentation, making the datalogical turn not only visible but central to identity co-production.

By recognizing identity co-production as both a response to and a product of the datalogical turn, we gain a more nuanced understanding of how digital spaces like TikTok function as environments where self-expression, datafication, and commodification intersect. The interplay of user agency and algorithmic bias thus becomes a defining characteristic of identity formation in the digital age, revealing that users' digital identities are not merely shaped by algorithms—they are actively co-produced within an ecosystem where data itself mediates selfhood. In this way, algorithmic identity co-production extends beyond individual expression to become a socio-technical process that reflects and reshapes cultural norms, reinforcing that our understanding of identity, agency, and social connection is now inseparable from the data systems that underlie our digital lives.

## **Conclusion**

In sum, these forms of engagement—tactical, experimental, and critical—not only illustrate identity co-production but also establish it as an interdisciplinary research area that demands attention across fields. The study of how users, particularly those with marginalized identities, interact with algorithms to shape their digital identities goes beyond the bounds of traditional media and

communication studies. It draws on insights from critical digital studies, sociology, psychology, and data ethics, reflecting the need to understand digital self-making within complex sociotechnical systems.

These users do not passively accept the content served to them; instead, they actively shape their digital identities by recognizing and responding to algorithmic biases. Through tactical engagement, they subtly navigate the algorithm's limitations, working within the system to influence its outputs. Experimental engagement amplifies this by allowing users to explore the algorithm's behaviors, using iterative tests to gain insights into its biases and refine their interactions. Critical engagement, though more passive, reflects an ongoing awareness that the algorithm mirrors broader societal biases—a recognition that shapes users' perceptions and behaviors on the platform.

### ***Limitations***

While this study addresses a gap in algorithmic identity research by offering insights into the perspectives and discursive practices of Black TikTok users, it has limitations in terms of external validity. The discursive practices observed are highly contextual, shaped by the specific sociocultural environment of TikTok, and rely on interpretive analysis, which introduces subjectivity. To mitigate this limitation, I have incorporated reflexive positionality throughout the study, continuously reflecting on my own influence and biases in interpreting the data. Furthermore, I grounded my analysis within established theories of algorithmic identity co-production to strengthen the rigor and consistency of my interpretations. A relevant limitation is that, while the concepts of user engagement—tactical, experimental, and critical—were identified through the experiences of Black users, they are theoretically applicable beyond this demographic. However, this study does not test their applicability to other user

groups, meaning further research is necessary to explore how users of different backgrounds experience and engage with algorithms.

The sample is limited by several constraints, particularly due to the lack of detailed information on how TikTok's default search algorithm filters content. Additionally, I was unable to restrict the scraping process to content posted within a specific timeframe, limiting control over the temporal scope of the dataset. The scraping tool itself was inconsistent, varying both in the volume and relevance of the videos it returned. Each scrape was conducted based on a specific search term, yielding relevant content but also including random, unrelated videos. The number of URLs returned per scrape fluctuated as well, making it challenging to maintain a consistent dataset.

To refine the dataset, I undertook a thorough cleaning process, reducing an initial collection of over 500 videos to a final sample of 138 videos. This approach, while imperfect, balanced the need for relevant content with the limitations of TikTok's search and scraping functionality, offering a method to collect data without manually selecting videos through the TikTok interface. Ultimately, this trade-off allowed for a more efficient collection of relevant content but introduced some limitations in sample consistency and temporal specificity.

Despite these limitations, the concepts presented here contribute theoretically to our understanding of user engagement with algorithms, particularly in the context of identity co-production. By establishing identity co-production as a distinct phenomenon rather than solely an analytical lens, this study opens space for future research to uncover new forms of user engagement. Centering on the experiences of Black users, this study underscores the importance of marginalized perspectives in

advancing algorithmic identity research and provides a foundation for future interdisciplinary studies to explore similar engagement patterns across diverse social and digital contexts.

### ***Future Research***

This study invites scholars across fields to consider the nuanced ways that users engage with consumer facing algorithms like the TikTok fyp. It highlights that identity co-production is not only about visibility or representation but involves complex negotiations between individuals and algorithmic systems, where critical digital literacy empowers users to navigate and reshape algorithmic influences on their identities. This expanded, interdisciplinary perspective advances our understanding of user-algorithm dynamics, establishing identity co-production as a multifaceted phenomenon that integrates theoretical approaches from a variety of disciplines to fully capture the digital self-making practices of marginalized users.

## WORKS CITED

- Abidin, C. (2021). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103.  
<https://doi.org/10.5334/csci.140>
- Banet-Weiser, S. (2012). *Authentic™: The Politics of Ambivalence in a Brand Culture*. New York University Press.
- Barta, K., & Andalibi, N. (2021). Constructing Authenticity on TikTok: Social Norms and Social Support on the "Fun" Platform. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–29. <https://doi.org/10.1145/3479574>
- Bhandari, A., & Bimo, S. (2022). Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society*, 8(1).
- Brock, A. (2016). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- Browne, S. (2015). *Dark Matters: On the Surveillance of Blackness*. Durham, NC: Duke University Press.
- Cheney-Lippold, J. (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, Culture & Society*, 28(6), 164-181.

Cotter, K. (2019). Playing the visibility game: How digital influencers and algorithms negotiate influence on Instagram. *New Media & Society*, 21(4), 895–913.

<https://doi.org/10.1177/1461444818815684>

DeVito, M. A., Birnholtz, J., & Hancock, J. T. (2017). Platforms, people, and perception: Using affordances to understand self-presentation on social media. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, 740-754.

DeVito, M. A. (2021). Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-38.

Eslami, M., Rickman, A., Vaccaro, K., et al. (2015). “I always assumed that I wasn’t really that close to [her]”: reasoning about invisible algorithms in news feeds. In *CHI ’15 Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 153–162. New York: ACM.

Clough, P. T., Gregory, K., Haber, B., & Scannell, R. J. (2015). The datalogical turn. In *Non-Representational Methodologies: Re-Envisioning Research* (pp. 146–164).

<https://doi.org/10.4324/9781315883540-13>Hargittai, E., & Hinnant, A. (2008). Digital inequality: Differences in young adults’ use of the Internet. *Communication Research*, 35(5), 602–621.

- Hearn, A. (2010). Structuring feeling: Web 2.0, online ranking and rating, and the digital 'reputation' economy. *Ephemera: Theory & Politics in Organization*, 10.
- Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic folk theories and identity: How TikTok users co-produce knowledge of identity and engage in algorithmic resistance. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 305, 1–44. <https://doi.org/10.1145/3476046>
- Lee, A. Y., Mieczkowski, H., Ellison, N. B., & Hancock, J. T. (2022). The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1-22.
- Papacharissi, Z. (2011). A networked self. *A networked self: Identity, community, and culture on social network sites*, 304-318.
- Peterson-Salahuddin, C. (2022). “pose”: Examining moments of “Digital” dark sousveillance on TikTok. *New Media & Society*.  
<https://doi.org/10.1177/146144482211080480>
- Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A., & Chavous, T. M. (1998). Multidimensional model of racial identity: A reconceptualization of African American racial identity. *Personality and Social Psychology Review*, 2(1), 18-39.
- Van Deursen, & Van Dijk (2014). Digital skills and the digital divide. *New Media & Society*, 16(3), 507–526.
- Volkova, K., & Kaun, A. (2019). Algorithmic Resistance: Social Media and Cultural Practices.



## **CHAPTER 3: EXPLORING ASSOCIATIONS BETWEEN PERCEIVED ALGORITHM RESPONSIVENESS, ALGORITHMIC AWARENESS, AND MEASURES OF RACIAL IDENTITY**

### **Abstract**

This study investigates how Black TikTok users perceive algorithmic responsiveness and insensitivity on the platform's For You Page (FYP), examining the roles of racial identity centrality, situational racial salience, and algorithm awareness. Addressing a critical gap in algorithm studies, this research centers on the experiences of a marginalized user group, building on the constructs of perceived algorithm responsiveness (PAR) and perceived algorithm insensitivity (PAI). Results indicate that users with higher racial identity centrality perceive the algorithm as both more responsive and more insensitive, suggesting that a strong racial identity amplifies both positive and negative algorithmic interactions. Situational racial salience positively predicts PAR but not PAI, highlighting how identity prominence enhances perceptions of responsiveness without necessarily influencing insensitivity. Additionally, awareness of human-algorithm interplay and content filtering positively predicts PAI, diverging from previous findings that linked awareness with reduced insensitivity. This discrepancy suggests that for Black TikTok users, heightened awareness of algorithmic personalization may intensify perceptions of insensitivity, possibly due to greater sensitivity to biases or identity misrepresentation. These findings expand the concept of algorithmic responsiveness within an identity co-production framework, suggesting that for Black TikTok users, greater awareness of algorithms can intensify perceptions of insensitivity, possibly due to heightened sensitivity to algorithmic biases or

misrepresentation of identity—an experience of being ‘filtered out’ by the algorithm that reinforces feelings of exclusion or misrecognition.

## **Introduction**

In an era where social media algorithms shape much of our online experiences, understanding how these systems interact with users' identities is increasingly essential, particularly for marginalized groups. While algorithms on platforms like TikTok's For You Page (FYP) personalize content based on user behaviors, they also have the potential to reinforce biases, amplify certain identities, and overlook others. Research on perceived algorithm responsiveness (PAR) and perceived algorithm insensitivity (PAI) sheds light on how users experience algorithms as either supportive or dismissive of their identities (Taylor & Choi, 2022). However, most studies on algorithmic responsiveness lack diversity in their samples, leaving critical gaps in our understanding of marginalized user experiences. This study addresses that gap by examining how Black TikTok users perceive the FYP algorithm's responsiveness and insensitivity in relation to their racial identity.

Building on identity co-production frameworks, this study explores how two dimensions of racial identity—centrality (the importance of racial identity to self-concept) and situational racial salience (the prominence of racial identity in specific contexts)—influence perceptions of algorithmic interactions. Additionally, I examine how dimensions of algorithm awareness contribute to perceptions of responsiveness, aiming to better understand the distinct but related roles of algorithm awareness and identity on user experience. By focusing on Black TikTok users, this research illuminates the complex ways in which algorithmic systems shape, validate, and potentially misrepresent marginalized identities in algorithmically curated environments.

## Literature review

### Perceived Algorithm Responsiveness

Taylor and Choi's (2022) research on perceived algorithm responsiveness bridges interpersonal communication theory with algorithm studies by introducing the concept of responsiveness to describe human-algorithm interaction. Traditionally, responsiveness, rooted in interpersonal interaction theory, has been essential for understanding intimacy development, where it involves an exchange of self-disclosure, validation, and understanding between partners (Reis & Shaver, 1988; Choi & Toma, 2022). Taylor and Choi repurpose this framework for algorithm studies, proposing that social media algorithms also engage users in a type of responsive relationship by curating content that either aligns with or misinterprets users' identities and goals. This concept adapts the four-step interpersonal responsiveness process—self-disclosure, enacted response, perceived responsiveness and insensitivity, and interaction outcomes—to the algorithmic domain, where users' online behaviors (e.g., likes, views) serve as "self-disclosures" to which algorithms respond with personalized content.

Taylor and Choi (2022) suggest that the concept of responsiveness can be applied to human–algorithm interactions, as people often carry over their interpersonal perceptions into interactions with machines, even those lacking human-like qualities (Gambino et al., 2020). In translating this concept to social media algorithms, Taylor and Choi argue for two dimensions of perceived algorithm responsiveness: Perceived Algorithm Responsiveness (PAR) and Perceived Algorithm Insensitivity (PAI). PAR reflects the degree to which users believe that algorithms understand, validate, and support their identity, as demonstrated by content that aligns with their goals and interests. In contrast, PAI captures perceptions that the algorithm misjudges or undermines users' identities, for instance, by ignoring

certain interests or failing to adapt to users' feedback. Notably, PAI is also designed to encompass biases that can marginalize certain identities, aligning with the broader recognition of algorithmic bias in decision-making (Noble, 2018).

This dual structure of PAR and PAI suggests a nuanced human-algorithm relationship where users may simultaneously feel understood by some aspects of the algorithm and overlooked or misrepresented by others. This coexistence of positive and negative perceptions echoes research showing that social media algorithms can help users find community but may also perpetuate stereotypes (Simpson & Semaan, 2020). The study thus conceptualizes responsiveness not as a static attribute of algorithms but as a dynamic, cyclical interaction that evolves over time, reflecting both positive and negative interactions with algorithmically curated content.

A core component of Taylor and Choi's framework is the role of algorithm awareness in enabling perceived responsiveness. Taylor and Choi's (2022) conceptualization of algorithm responsiveness builds on the foundation of algorithm awareness. Given the hidden nature of algorithms, research has shown that awareness of their operation is essential before users can form perceptions about them (Eslami et al., 2015; Shin, 2020). They argue that awareness—particularly content filtering and human-algorithm interplay—is a prerequisite for perceiving responsiveness but remains distinct from PAR and PAI. Content filtering awareness reflects users' understanding that social media content is personalized through algorithms, while human-algorithm interplay awareness involves a more advanced understanding of how user data is employed to tailor content (Zarouli et al., 2021). Although awareness does not equate to perceptions of responsiveness, it enables users to interpret algorithmic actions as either aligned with or insensitive to their identity, thus facilitating the responsiveness process.

By integrating interpersonal responsiveness with algorithm awareness, Taylor and Choi propose that user engagement with social media algorithms mirrors interpersonal relationships in important ways, such as influencing self-concept and user satisfaction. The study suggests that perceived responsiveness—whether supportive (PAR) or insensitive (PAI)—influences users' overall engagement and satisfaction, which may extend to feelings of belonging or marginalization based on algorithmic biases. Taylor and Choi's work ultimately positions algorithm responsiveness as a critical framework for understanding how identity and algorithmic curation intersect in complex ways, emphasizing the importance of conceptualizing human-algorithm interaction as a nuanced, ongoing relationship shaped by both user behavior and algorithmic design.

### **Algorithm Identity Co-production**

Early theories on self-representation in visual media and online networks often overlook the influence of algorithmic power on self-presentation. As a result, the established 'networked self' framework has evolved into the concept of the 'algorithmized self' to more accurately encompass the complexities of these algorithm-driven interactions. While the 'networked self' emphasizes self-construction through a "reflexive process of fluid associations with social circles" (Papacharissi, 2011), the 'algorithmized self' shifts this focus toward identity co-production via "reflexive engagement with previous self-representations rather than with one's social connections" (Bhandari & Bimbo, 2021). This framework introduces a self-reflective user engagement wherein users continuously interact with algorithm-mediated versions of themselves. This repeated interaction with algorithmically mediated self-representations enhances users' capacity to shape their identity, while simultaneously being influenced by algorithmic outputs. Consequently, examining this dual engagement—where users both

shape and are shaped by digital representations—positions TikTok users uniquely in navigating the algorithm's external influence alongside their internal self-reflection.

While the 'algorithmized self' concept emphasizes user agency through the co-production of identity, scholars focusing on folk theories are more concerned with systemic biases and how users co-produce knowledge of identity through their interactions with the FYP algorithm. For instance, Karizat et al. (2021) introduce the Identity Strainer Theory to capture the belief that the FYP algorithm 'filters out' and suppresses marginalized identities based on factors such as race, body size, ability, class, LGBTQ status, and political affiliations. Additionally, the authors introduce the concept of 'algorithmic privilege' to describe the advantage experienced by users whose identities align with the algorithm's biases, in contrast to 'algorithmic representational harm,' which describes the negative experiences faced by users without this privilege. Furthermore, the authors interpret users' efforts to modify the FYP as a form of 'algorithmic resistance' (Volkova & Kaun, 2019), aimed at 'repairing' the algorithm, thus underscoring the agency users exercise in navigating algorithmically curated environments.

Taylor and Choi (2022), who developed the perceived algorithm responsiveness measure, refer directly to the co-productionist approach in their literature review:

Karizat and colleagues (2021) see the relationship between algorithms and identity as a co-production, where people's identity influences algorithmic curation but, in turn, algorithmic curation influences the identity development of the user. This co-production prioritizes some facets of a person's self-concept over others and exposes marginalized groups to algorithmic harms. However, people tend to hold more positive attitudes toward social media algorithms when the personalized content represents the multifaceted, dynamic aspects of the self (Lee et

al., 2022). We build on this work at the intersection of algorithms and identity by repositioning this co-production as human–algorithm interaction. This theoretical invention allows for interpersonal communication theory to elaborate on the processes experienced during human–algorithm interaction. (p.2)

The authors reframe co-production using the algorithmic crystal theory (Lee et al, 2022) to move away from a more critical perspective such as identity strainer theory. ‘Algorithmic crystal’ refers to the way algorithms interact with user identity by reflecting self-concepts that are both multifaceted and dynamic. Moreover, they shape perspectives on others encountered through the algorithm, orienting users to recognize parts of themselves ‘refracted’ in other users and to experience ephemeral, diffracted connections with groups of similar others.

By reinterpreting co-production as human-algorithm interaction, however, the authors risk recoding user perceptions of algorithmic bias as mere "insensitivity." This reframing inherently shifts theoretical discussions towards user agency and subjective experiences while overlooking pre-existing algorithmic structures. Consequently, this approach may obscure deeper understandings of how systemic racial bias—whether real or perceived—manifests in the FYP user experience. This oversight can lead to inadvertent dismissal of how machine learning models and datasets embed and perpetuate inequalities. While user perceptions are indeed central to identity co-production, they operate within a framework shaped by algorithmic structures, making it essential to integrate critical perspectives on these systemic biases into frameworks like algorithmic crystal and algorithmic responsiveness. Human-algorithm interaction provides a valuable lens for empirically measuring user perceptions, yet identity

co-production better describes the broader, ongoing interaction between user identity and the FYP algorithm.

Concepts associated with TikTok's personalization, such as the algorithmic crystal (Lee et al., 2022), meaningfully contribute to the co-production of algorithmic identity. While Taylor and Choi describe human-algorithm interaction as akin to co-production, framing perceived algorithm responsiveness within a more expansive co-production model allows for a deeper understanding of the complex interdependencies between user identity and the FYP algorithm. This study underscores that viewing human-algorithm interaction through the critical perspective of identity co-production highlights how algorithms subtly shape and reflect user identities. This approach not only reveals how the algorithm dynamically mirrors aspects of the self but also exposes broader systemic biases embedded in its design, capturing a fuller picture of algorithmic influence on user identity.

Through this lens, identity co-production emerges as an iterative process wherein marginalized users perceive encoded bias in the fyp algorithm. This approach expands our understanding of digital identity in the algorithmic age, providing a more nuanced view of how Black users' interactions with the FYP become both acts of self-representation and sites of subtle resistance to algorithmic bias. By examining the relationship between measures of racial identity and perceived algorithm responsiveness, this study offers a critical perspective on how algorithms shape, influence, and are influenced by the identities they interact with, making the co-production framework a valuable tool for exploring the complexities of identity in algorithmically mediated spaces.

### **Racial Identity Centrality and Salience**



Racial identity centrality is a key concept from the Multidimensional Inventory of Black Identity (MIBI) and the Multidimensional Model of Racial Identity (MMRI) (Sellers et al., 1998). Centrality refers to how much individuals define themselves regarding their racial group. It represents the importance of race to an individual's self-concept and is a crucial dimension in understanding the complexity of Black identity. Within the MMRI framework, racial identity is not seen as a singular or static experience but rather as a multifaceted construct that can vary across situations and over time. Centrality is one of the four dimensions of racial identity in this model, highlighting how integral racial identity is to an individual's overall sense of self.

Racial Identity Salience, in contrast, is a more dynamic aspect of racial identity that fluctuates depending on the situational context. Unlike centrality, which is relatively stable, racial salience varies based on how relevant racial identity feels in a given moment or setting (Douglass & Umana-Taylor, 2015). Within the MMRI framework, salience provides a lens for understanding how the significance of racial identity shifts in response to environmental cues, such as being in a racially homogenous or heterogeneous space. This fluidity allows individuals to experience their racial identity differently across contexts, with salience becoming heightened in situations where race is a prominent factor, such as in discussions or environments where racial issues are explicitly addressed.

In digital spaces, especially those mediated by algorithms, both centrality and salience may play unique roles in shaping user perceptions of algorithmic responsiveness. Racial identity centrality might lead users to consistently interpret algorithmic interactions through a racialized lens, reinforcing the significance of race in digital interactions. In contrast, racial salience may prompt situationally heightened awareness of race, potentially altering how users respond to specific algorithmic content on

platforms like the For You Page (FYP) of TikTok. Together, centrality and salience provide a nuanced view of racial identity, capturing both the stable significance of race in self-concept and the situational relevance of race in user experiences, making them critical for examining how marginalized identities are represented, supported, or misrepresented by algorithms. This duality within the MMRI framework is instrumental in exploring the complexities of racial identity within algorithm-driven platforms.

### **Initial Hypotheses**

#### **H1: Centrality will be positively associated with racial salience on the FYP.**

In the context of TikTok's For You Page (FYP), individuals with higher racial centrality may be more attuned to algorithmic representations of race, which could heighten the salience of their racial identity. This aligns with prior research suggesting that those who place a strong emphasis on racial identity are more likely to interpret and respond to social interactions through a racialized lens (Sellers et al., 1998; Douglass & Umana-Taylor, 2015). On platforms like TikTok, where algorithmic curation can amplify or suppress certain identities, this relationship may be particularly pronounced, as users navigate racialized content curation in real time.

#### **H2: Higher levels of awareness (total AMCA score) will be negatively associated with total perceived responsiveness (PAR and PAI combined score).**

This expectation was based on Taylor and Choi's (2022) finding that PAI has small negative associations with both awareness of content filtering and human-algorithm interplay, indicating that as users become more aware of algorithmic personalization, their perceptions of insensitivity may slightly diminish. In other words, increased awareness of personalization mechanisms may lead users to recognize biases or errors in algorithmic curation, potentially heightening perceptions of insensitivity

(PAI). Conversely, greater awareness might slightly reduce perceptions of insensitivity as users attribute algorithmic actions to system design rather than personal targeting. However, the combined score of PAR and PAI reflects both positive and negative perceptions, and this duality makes it likely that heightened awareness creates a net negative association with total perceived responsiveness. By examining this relationship, the study seeks to clarify the nuanced role of awareness in shaping user experiences with algorithms.

**H3: Higher levels of racial salience on the For You Page (FYP) will be associated with greater perceived algorithm insensitivity (PAI) but not necessarily with perceived algorithm responsiveness (PAR).**

This hypothesis stems from research suggesting that situational racial salience intensifies users' sensitivity to racialized interactions, including algorithmic curation. Although in offline social contexts racial salience is a positive experience in some contexts (Douglass et. al., 2016), when racial identity becomes more salient in an algorithmic context—such as when users encounter racially coded content or perceive exclusion of their racial identity—they may interpret these interactions as algorithmic insensitivity (Douglass & Umana-Taylor, 2015; Karizat et al., 2021). In contrast, perceived algorithm responsiveness (PAR) may not show a similar relationship because responsiveness requires not only the absence of insensitivity but also active validation and support of the user's identity. Thus, while racial salience might heighten users' recognition of algorithmic failures (PAI), it may not guarantee experiences of responsiveness (PAR), highlighting the asymmetric relationship between these constructs in algorithmic environments.

## **Methods**

## **Sampling, Recruiting, and Data Collection**

Participants were recruited through Prolific, a research platform connecting researchers to participants based on screener surveys. I designed a screener to select U.S.-based participants who identify as Black/African American and are active TikTok users, defined by Prolific as using the platform at least once per month. The study information available to participants emphasized the focus on the TikTok algorithm, aiming to engage those familiar with interacting on the For You Page (FYP). No restrictions were placed on age or gender, as there are no existing theoretical reasons to target specific cohorts. Given that a study like this had not been conducted before, I based my sample size on Taylor and Choi's (2022) approach, which used 406 participants to detect a small effect ( $f^2 = .05$ ) with at least 80% power. This provided a benchmark for achieving reliable statistical power in a study exploring nuanced perceptions of algorithm responsiveness and identity, especially within a marginalized user group.

After passing the screener, participants accessed the survey through an external Qualtrics link. To ensure privacy, the survey did not collect personal information; instead, participants provided their Prolific ID after consenting. Demographic data were provided by Prolific and later merged with survey responses using matching IDs. I received 507 responses, with participants compensated based on survey completion time (average time: 11.82 minutes), and additional funds distributed as bonuses.

Following data cleaning in Qualtrics, which included removing low-quality responses (e.g., straight-lining, incompletes) and excluding those who did not identify as Black/African American, 445 eligible responses were retained for analysis. The final sample included 61.3% females, 38.4% males, and 0.2% who preferred not to disclose gender, with an average age of ( $SD = 11.39$ ). The median age

was 37, with 25% of participants aged 29 or younger and another 25% aged 49 or older, indicating a diverse age distribution.

### **Survey Measures**

All measures were rated on a 5-point Likert scale (1 = not at all, 5 = completely).

***Perceived Algorithm Responsiveness (PAR and PAI).*** PAR measures users' perceptions of algorithmic understanding, acknowledgment, and support for their identities, with higher PAR scores indicating alignment between user identity and feed content. The PAR scale includes 8 items (e.g., "The TikTok algorithm understands me"), ( $\alpha = .91$ ).

PAI measures perceived algorithmic misjudgment or restriction of user identity and goals. Higher PAI scores indicate that the algorithmic content diverges from the user's purpose on the platform. The PAI scale includes 7 items (e.g., "The [TikTok] algorithm dismisses my interests";  $\alpha = .93$ ). A two-factor CFA model for the PAR and PAI items confirmed their distinctiveness, with a scree plot and eigenvalues indicating these factors captured significant variance.

***Algorithmic Media Content Awareness (AMCA).*** The AMCA Scale (Zarouali et al., 2021) assesses user awareness across four dimensions: content filtering ( $\alpha = .82$ ), automated decision-making ( $\alpha = .74$ ), human-algorithm interplay ( $\alpha = .77$ ), and ethical considerations ( $\alpha = .57$ ; Spearman-Brown coefficient = 0.73). A CFA confirmed the expected four-factor structure.

***Racial Identity Centrality.*** Racial centrality measures the importance of racial identity within one's self-concept. (Sellers et. al., 1998) I analyzed this construct using an eight-item scale (e.g., "In general, being Black is an important part of my self-image"). The factor analysis showed a dominant primary factor

(Eigenvalue = 3.94) and a smaller secondary factor (Eigenvalue = 1.15), however factor loadings were appropriate to move forward with the one factor structure ( $\alpha = .83$ ).

**Racial Identity Salience on the FYP.** Racial salience is context-dependent, so I adapted four items from Worrell et al. (2020) to measure racial salience in the TikTok FYP context (e.g., “When I go on my FYP, I always take note of the racial make-up of the creators in my feed”;  $\alpha = .70$ ). A CFA confirmed one-dimensionality, with Factor 1 being the only factor with an eigenvalue over 1 (Eigenvalue = 2.11), and item loadings between 0.54 and 0.67 indicating strong internal consistency.

## Regression Models

**H1: Centrality will be positively associated with racial salience on the FYP.**

$$Y_{\text{Salience}} = \beta_0 + \beta_{\text{Centrality}} \cdot x_{\text{Centrality}} + \epsilon \quad (1)$$

To begin my exploratory analysis, only a simple regression model is necessary to address my first hypothesis, which essentially asks if race is relevant to users in their FYP experience. The more important Black identity is to users’ self-concept, the more relevant it will be to their experiences with the algorithm. This model provides context for my adapted salience measure. I isolate the racial identity measures to test if the data matched the theory, at least to the extent that racial identity centrality and racial salience on the FYP were positively associated.

**H2: Higher levels of awareness (total AMCA score) will be negatively associated with total perceived responsiveness (PAR and PAI combined score).**

(2)

$$Y_{RSP} = \beta_0 + \beta_{\text{Salience}} \cdot x_{\text{Salience}} + \beta_{\text{Centrality}} \cdot x_{\text{Centrality}} + \beta_{\text{Awareness}} \cdot x_{\text{Awareness}} + \epsilon$$

Multiple regression models are powerful tools because they allow us to estimate the impact of<sup>(2)</sup> one variable on a dependent variable while accounting for the influence of other explanatory variables in the model. This capability is crucial when the explanatory variables are correlated, as it enables a more accurate understanding of how each variable independently relates to the outcome. Even when the explanatory variables are uncorrelated, or orthogonal, including them in the model can help reduce unexplained variance, leading to more precise estimates of the relationships being studied.

**H3: Higher levels of racial salience on the For You Page (FYP) will be associated with greater perceived algorithm insensitivity (PAI) but not necessarily with perceived algorithm responsiveness (PAR).**

$$Y_{\text{PAI}} = \beta_0 + \beta_{\text{Salience}} \cdot x_{\text{Salience}} + \beta_{\text{Centrality}} \cdot x_{\text{Centrality}} + \beta_{\text{Awareness}} \cdot x_{\text{Awareness}} + \epsilon \quad (3)$$

$$Y_{\text{PAR}} = \beta_0 + \beta_{\text{Salience}} \cdot x_{\text{Salience}} + \beta_{\text{Centrality}} \cdot x_{\text{Centrality}} + \beta_{\text{Awareness}} \cdot x_{\text{Awareness}} + \epsilon \quad (4)$$

This hypothesis calls for two separate regression models, one with PAI as the dependent variable and the other with PAR as the dependent variable. Now that perceptions of the algorithm responsiveness are the outcome variable, I included the other measures in these models to control racial identity centrality and algorithm awareness.

In addition to the planned analyses, several exploratory tests were conducted to investigate emerging patterns observed during the initial data analysis. These included tests for mediation effects of awareness, specifically focusing on the human-algorithm interplay dimension. These additional analyses were deemed necessary to provide a more comprehensive understanding of the relationships between racial identity centrality and perceived algorithm responsiveness.

For example, I also included a model with an interaction term  $x_{\text{Salience}} \cdot x_{\text{Centrality}}$  to test whether the effect of racial salience on perceived algorithm responsiveness changes at different levels of centrality. The interaction term changes the interpretation of effects ( $\beta$ ) of racial salience and centrality, as they only apply when the other interacting variable is zero. Therefore, if the interaction term is significant, to understand the full effect of salience on responsiveness at any given level of centrality requires the main effect and the interaction effect. I also ran more granular models that included the individual dimensions of awareness and both PAR and PAI as outcome variables to build on learnings from my initial hypotheses. Each additional model or test throughout my exploratory analysis is accompanied by the theoretical and mathematical explanation throughout the results section.

### Exploratory Analysis

**Multicollinearity.** None of the variables showed strong correlations, although there were moderate correlations between the human-algorithm interplay, automated decisions, and content filtering dimensions of awareness (Table 1) as expected for subscales of the same measure. To further confirm that multicollinearity was not an issue, I performed a variance inflation factor (VIF) analysis. All VIF values were below the threshold of 5, with content filtering having the highest value at 1.7, indicating that multicollinearity is minimal, and the regression coefficients are reliable (Table 2).

	<i>M</i>	<i>SD</i>	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>
1. Centrality	31.24	6.54	-					
2. Salience	12.03	4.18	0.30**	-				



<i>3. Content Filtering</i>	16.34	3.53	0.04	0.11*	-			
<i>4. H/A Interplay</i>	11.64	2.97	0.02	0.12**	0.54**	-		
<i>5. Automated Decisions</i>	10.64	3.06	-0.05	0.15**	0.54**	0.50**	-	
<i>6. Ethical Concerns</i>	9.34	2.96	-0.10*	0.18**	0.34**	0.26**	0.31**	-

Table 1: Correlation Matrix

\* $p < .05$ . \*\* $p < .01$ .

Table 2: VIF Values

<b>Variable</b>	<b>VIF</b>
<i>Centrality</i>	1.15
<i>Salience</i>	1.17
<i>Content Filtering</i>	1.70
<i>H/A Interplay</i>	1.56
<i>Automated Decisions</i>	1.60
<i>Ethical Concerns</i>	1.22

**Normality of Residuals.** I assessed the normality of the residuals of the model including all variables, including the individual awareness dimensions, using a Q-Q plot (Figure 1), which indicated a moderate left skew, particularly at the lower tail. This deviation from normality was confirmed by the Shapiro-Wilk test ( $p < 0.001$ ), suggesting non-normality. Additional indicators included a skewness of -0.82 and a

kurtosis of 3.47, slightly above the normal value of 3, as well as a Jarque-Bera test result ( $p = 1.92 \times 10^{-12}$ ), which suggested heavier tails than a normal distribution. However, these deviations were minor, and the Durbin-Watson statistic of 2.08 indicated no autocorrelation in the residuals. The slight skew and kurtosis are not expected to affect the robustness of the regression analysis.

Figure 1: Q-Q Plot of Residuals for model with all predictor variables (ind. AMCA dimensions)

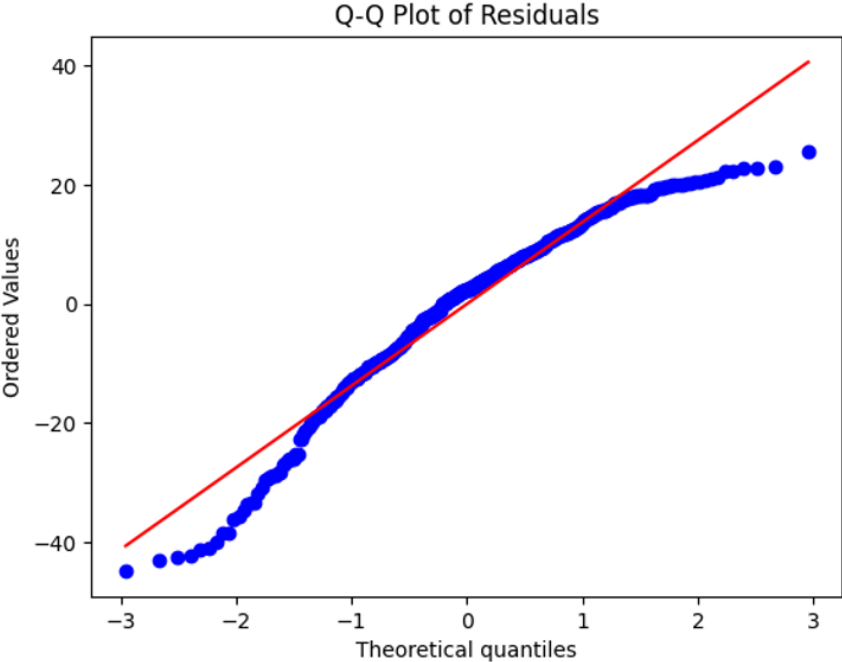
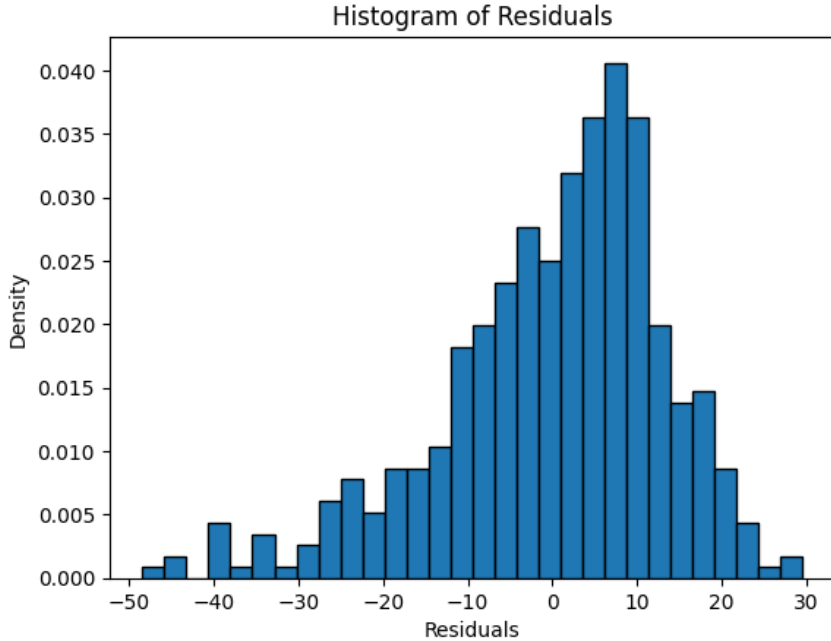


Figure 2: histogram of exhaustive model residuals



**Outlier Assessment.** I checked for outliers using Z-scores, (see Appendix C). This revealed a only 2 outliers, but they were unlikely to affect regression results. As such, no transformations or non-parametric methods were applied. Additionally, low-quality responses were removed during data collection in Qualtrics, resulting in no missing values across the primary variables.

### **Initial Hypotheses and Exploratory Analyses**

I began this study with a set of exploratory hypotheses to guide my investigation into how racial identity centrality, racial identity salience, and dimensions of algorithm awareness (such as content filtering and human-algorithm interplay) impact Black TikTok users' perceptions of algorithm responsiveness and insensitivity. Given the complexity and evolving nature of human-algorithm interactions, especially among marginalized groups, the exploratory approach was crucial to uncovering novel insights without preconceived assumptions. By beginning with broad hypotheses rather than

narrowly defined predictions, I aimed to capture the nuanced ways these identity dimensions might influence users' perceptions of the algorithm's responsiveness to their identity and motivations. This exploratory framework allowed for a deeper investigation into how these concepts interact within algorithm-driven environments, particularly in understudied spaces like TikTok's For You Page (FYP).

#### *H1: Racial Salience and Racial Centrality*

I first tested the relationship between racial centrality and salience, predicting that users with high racial centrality will also have high racial salience. Regression results showed that centrality significantly predicted salience ( $\beta = 0.30$ ,  $p < 0.001$ ,  $R^2 = 0.09$ ). This also provided context for the external validity of my adapted racial salience scale.

#### *H2: Awareness and Total Perceived Responsiveness*

Hypothesis 2 proposed that greater awareness would be associated with decreased perceived responsiveness. This expectation was based on Taylor and Choi's (2022) finding that PAI has small negative correlations with both awareness of content filtering and human-algorithm interplay, suggesting that as users become more aware of algorithmic personalization, their perceptions of insensitivity may slightly diminish. However, the results did not support this hypothesis (Table 3).

Contrary to expectations, awareness (measured as AMCA) significantly predicted perceived responsiveness ( $\beta = 1.60$ ,  $p = 0.02$ ), with each standard deviation increase in awareness linked to a 1.60 standard deviation increase in responsiveness. This finding suggests that users with higher algorithmic awareness perceive the algorithm as more responsive to their identities, contrary to prior assumptions.

This result may reflect a deeper understanding of algorithmic personalization, where increased knowledge of how the system operates fosters a sense of control or engagement with the platform.

In this model, racial identity centrality also emerged as a strong predictor of perceived responsiveness ( $\beta = 2.90$ ,  $p < 0.001$ ), indicating that users who view their Black identity as central to their self-concept perceive the algorithm as significantly more responsive. This finding aligns with existing theories on self-concept and user engagement, suggesting that users with strong identity centrality may recognize more cues of responsiveness in algorithmic outputs.

Racial salience, on the other hand, had a small, non-significant association with responsiveness ( $\beta = 0.12$ ,  $p = 0.87$ ). This result suggests that salience, or the situational prominence of one's racial identity, does not meaningfully predict how users perceive the responsiveness of the algorithm. The lack of significance here supports the idea that while racial identity centrality has a stable influence on perceived responsiveness, momentary or situational shifts in identity salience may not play a substantial role.

These results highlight the distinct roles of centrality and awareness in shaping perceptions of algorithmic responsiveness. While centrality operates as a stable self-concept predictor, awareness appears to promote an active sense of engagement, both of which contribute to greater perceived responsiveness. The non-significant effect of salience underscores the need to further explore how

stable identity traits differ from more fluid, context-dependent identity shifts in influencing user-algorithm interactions.

*Table 3: Total Awareness and Total Responsiveness Regression Results*

	<b>Coefficient</b>	<b>SE</b>	<b>95% CI</b>		<b>p</b>
			<b>Lower Limit</b>	<b>Upper Limit</b>	
<i>Constant</i>	47.21	0.67	45.90	48.52	0.00
<i>Centrality</i>	2.90	0.70	1.52	4.23	0.00
<i>Salience</i>	0.12	0.71	-1.29	1.52	0.87
<i>AMCA Total</i>	1.60	0.68	0.27	2.94	0.02

*H3: Racial Salience and Perceived Responsiveness Factors (PAR and PAI)*

Although the two-factor structure is designed to be one measure, Taylor and Choi (2022) do not provide any comparative context for total perceived responsiveness and perform their analysis examining PAR and PAI separately. Furthermore, I performed further analysis at the dimension level to unpack the relationship between predictor variables and perceived responsiveness.

In the models including total awareness and centrality, racial salience was not a significant predictor of PAR or PAI. For PAR, the coefficient for salience was  $\beta = 0.51$  ( $p = 0.19$ ), and for PAI, the coefficient for salience was  $\beta = -0.40$  ( $p = 0.33$ ). These results indicate that racial salience does not significantly predict perceptions of either total algorithm responsiveness (PAR) or insensitivity (PAI) when controlling for awareness and centrality. The findings highlight that other factors, such as

awareness and centrality, play a more influential role in shaping these perceptions. This reinforces the notion that stable identity traits (like centrality) have a stronger influence on user-algorithm interactions than situational or context-dependent features (like salience).

Table 4: PAR and AMCA Total Regression Results

	<b>Coefficient</b>	<b>SE</b>	<b>95% CI</b>		<b>p</b>
			<b>Lower Limit</b>	<b>Upper Limit</b>	
<i>Constant</i>	22.81	0.37	22.01	23.53	.00
<i>Centrality</i>	1.39	0.39	0.63	2.15	.00
<i>Salience</i>	0.51	0.39	-0.26	1.28	.19
<i>AMCA Total</i>	0.75	0.38	0.02	1.49	.05

Table 5: PAI and AMCA Total Regression Results

	<b>Coefficient</b>	<b>SE</b>	<b>95% CI</b>		<b>p</b>
			<b>Lower Limit</b>	<b>Upper Limit</b>	
<i>Constant</i>	24.40	0.38	23.66	25.14	.00
<i>Centrality</i>	1.51	0.40	0.74	2.29	.00
<i>Salience</i>	-0.40	0.40	-1.19	0.40	.33
<i>AMCA Total</i>	0.85	0.38	0.01	1.60	.03

### AMCA Scale Dimensions

To better understand the relationship between the variables, I conducted additional analyses using each AMCA dimension individually. This approach increased the model's adjusted R-squared from 0.05 to 0.13, indicating that the individual dimensions capture more variance than the total awareness score.

#### *Total Responsiveness as Outcome*

In the regression model predicting total responsiveness, I included the four AMCA dimensions—content filtering, human-algorithm interplay, automated decision-making, and ethical concerns—along with centrality and racial salience (Table 6). Centrality emerged as a significant positive predictor ( $\beta = 2.18, p = 0.001$ ), indicating that users who view their racial identity as central to their self-concept perceive the algorithm as more responsive.

Human-algorithm interplay was the strongest positive predictor of responsiveness ( $\beta = 3.81, p < 0.001$ ), suggesting that users who engage directly with the algorithm perceive it as more responsive. Content filtering also had a significant positive association with responsiveness ( $\beta = 1.64, p = 0.048$ ), indicating that users who recognize content filtering processes perceive the algorithm as more responsive.

In contrast, ethical concerns had a significant negative impact on responsiveness ( $\beta = -3.20, p < 0.001$ ), indicating that users with greater ethical concerns about the algorithm view it as less responsive.



Neither racial salience ( $\beta = 0.72$ ,  $p = 0.298$ ) nor automated decision-making ( $\beta = -1.01$ ,  $p = 0.211$ ) were significant predictors of responsiveness, suggesting they play a lesser role in shaping users' perceptions of algorithmic responsiveness.

These findings highlight that specific dimensions of algorithmic awareness—particularly human-algorithm interplay, content filtering, and ethical concerns—along with centrality, are key factors in shaping users' perceptions of algorithm responsiveness. The results suggest that users who recognize how their actions influence the algorithm, understand content filtering, or see the system as ethically questionable have stronger perceptions of responsiveness—whether positive or negative—depending on the nature of their engagement.

Table 6: Total Responsiveness with AMCA Dimensions Regression Results

	<b>Coefficient</b>	<b>SE</b>	<b>95% CI</b>		<b>p</b>
			<b>Lower Limit</b>	<b>Upper Limit</b>	
<i>Constant</i>	47.21	0.64	45.96	48.46	.00
<i>Centrality</i>	2.18	0.68	0.84	3.52	.00
<i>Salience</i>	0.72	0.69	-0.64	2.07	.30
<i>Content Filtering</i>	1.64	0.83	0.01	3.28	.05
<i>H/A</i>	3.81	0.79	2.25	5.37	.00
<i>Interplay</i>					
<i>Automated Decisions</i>	-1.00	0.81	-2.6	0.58	.21
<i>Ethical Concerns</i>	-3.20	0.70	-4.58	-1.82	.00

### PAR and PAI as Separate Outcomes

To gain a more nuanced view, I analyzed PAR (Perceived Algorithm Responsiveness) and PAI (Perceived Algorithm Insensitivity) separately. This approach clarified how each AMCA dimension influences perceptions of responsiveness and insensitivity.

Table 7: PAR as Outcome with AMCA Dimensions Regression Results

	<i>Coefficient</i>	<i>SE</i>	<i>95% CI</i>		<i>p</i>
			<i>Lower Limit</i>	<i>Upper Limit</i>	
<i>Constant</i>	22.81	0.36	22.11	23.51	.00
<i>Centrality</i>	1.09	0.38	0.34	1.84	.00
<i>Saliency</i>	0.76	0.39	0.00	1.52	.05
<i>Content Filtering</i>	0.42	0.47	-0.50	1.33	.37
<i>H/A Interplay</i>	1.92	0.45	1.05	2.80	.00
<i>Automated Decisions</i>	-0.33	0.45	-1.22	0.55	.46
<i>Ethical Concerns</i>	-1.36	0.39	-2.13	-0.59	.00

In the regression model with PAR as the outcome (Table 7), several key predictors emerged as significant influences on perceptions of algorithm responsiveness. Centrality was a significant positive predictor ( $\beta = 1.09$ ,  $p = 0.004$ ), indicating that users who consider their racial identity central to their self-concept are more likely to perceive the algorithm as responsive to their identity.

Racial salience showed a marginally significant positive relationship with PAR ( $\beta = 0.76, p = 0.050$ ), suggesting that users with higher racial salience may also be more likely to feel that the algorithm is responsive to their identity. The strongest positive predictor of PAR was human-algorithm interplay ( $\beta = 1.92, p < 0.001$ ), indicating that users who actively engage with the algorithm perceive it as more responsive. This finding highlights the importance of direct user interaction in shaping perceptions of algorithmic responsiveness.

Ethical concerns had a significant negative impact on PAR ( $\beta = -1.36, p = 0.001$ ), suggesting that users with greater ethical concerns about the algorithm are less likely to perceive it as responsive. Neither content filtering ( $\beta = 0.42, p = 0.370$ ) nor automated decisions ( $\beta = -0.33, p = 0.461$ ) were significant predictors of PAR, indicating they play a lesser role in shaping perceptions of algorithm responsiveness.

These findings highlight centrality, racial salience, and particularly human-algorithm interplay as key factors in shaping user perceptions of algorithm responsiveness, while ethical concerns diminish these perceptions. The results suggest that stable identity traits (like centrality), situational identity shifts (like salience), and active user engagement (like interplay) are central to understanding how users perceive algorithmic responsiveness.

*Table 8: PAI as Outcome with AMCA Dimensions Regression Results*

<i>Coefficient</i>	<i>SE</i>	<i>95% CI</i>		<i>p</i>
		<i>Lower Limit</i>	<i>Upper Limit</i>	

<i>Constant</i>	24.40	0.36	23.69	25.11	.00
<i>Centrality</i>	1.09	0.39	0.33	1.84	.00
<i>Salience</i>	-0.04	0.39	-0.80	0.73	.92
<i>Content Filtering</i>	1.23	0.47	0.30	2.15	.00
<i>H/A Interplay</i>	1.89	0.45	1.00	2.77	.00
<i>Automated Decisions</i>	-0.68	0.46	-1.57	0.22	.14
<i>Ethical Concerns</i>	-1.84	0.40	-2.62	-1.06	.00

In the regression model with PAI as the outcome (Table 8), several predictors emerged as significant influences on perceptions of algorithmic insensitivity.

Centrality was a significant positive predictor ( $\beta = 1.09$ ,  $p = 0.01$ ), indicating that users who consider their racial identity as central to their self-concept are more likely to perceive the algorithm as insensitive.

Content filtering also had a significant positive effect on insensitivity ( $\beta = 1.23$ ,  $p = 0.01$ ), suggesting that users with greater awareness of content filtering are more likely to perceive the algorithm as insensitive.

The strongest positive predictor of insensitivity was human-algorithm interplay ( $\beta = 1.89$ ,  $p < 0.01$ ), indicating that users who actively engage with the algorithm tend to perceive it as more insensitive. This finding underscores the impact of user engagement on perceptions of algorithmic fairness and responsiveness.

Ethical concerns had a significant negative impact on PAI ( $\beta = -1.84, p < 0.01$ ), suggesting that users with stronger ethical concerns about the algorithm perceive it as less insensitive. This may reflect a more critical stance toward the algorithm's ethical implications, leading to lower perceptions of insensitivity.

Neither racial salience ( $\beta = -0.04, p = 0.92$ ) nor automated decisions ( $\beta = -0.68, p = 0.14$ ) were significant predictors of PAI, indicating they do not play a meaningful role in shaping perceptions of algorithmic insensitivity in this model.

These findings highlight centrality, content filtering, and human-algorithm interplay as key drivers of perceptions of algorithmic insensitivity, while ethical concerns reduce these perceptions. This suggests that stable identity traits (like centrality), user engagement (like interplay), and awareness of content filtering processes are critical in shaping perceptions of algorithmic insensitivity.

### **Racial Centrality and Racial Salience Moderation Analysis**

I conducted a moderation analysis to examine whether the effect of racial identity salience on perceived algorithm responsiveness (PAR) depends on racial identity centrality. This analysis is grounded in the idea that the stable importance of racial identity within an individual's self-concept (centrality) may influence how relevant that identity feels in specific contexts, such as on social media platforms like TikTok (salience).

Centrality reflects the stable significance of racial identity within an individual's broader self-concept, while salience captures how prominently racial identity is perceived in a specific context, like TikTok's For You Page (FYP). I hypothesized that centrality might amplify the effect of salience on perceptions of algorithmic responsiveness. In other words, for individuals with high centrality, an

increase in salience on the FYP might have a stronger impact on how responsive they perceive the algorithm to be. This moderation analysis tested whether users who view their racial identity as central to their self-concept are more sensitive to its relevance on the platform, potentially shaping their experience of algorithmic responsiveness.

In the analysis (Table 9), centrality was a significant positive predictor of PAR ( $\beta = 1.25$ ,  $p = 0.001$ ), indicating that users with higher racial identity centrality perceive the algorithm as more responsive. Human-algorithm interplay also had a significant positive effect ( $\beta = 1.95$ ,  $p < 0.001$ ), suggesting that users who actively engage with the algorithm tend to perceive it as more responsive.

Ethical concerns had a significant negative impact on PAR ( $\beta = -1.36$ ,  $p = 0.001$ ), indicating that users with greater ethical concerns about the algorithm perceive it as less responsive.

Salience had a marginally significant positive relationship with PAR ( $\beta = 0.71$ ,  $p = 0.07$ ), suggesting that users who see their racial identity as relevant on the FYP may view the algorithm as more responsive, although this result was not statistically significant at the 0.05 level.

The interaction term between centrality and salience was also not statistically significant ( $\beta = 0.64$ ,  $p = 0.07$ ), indicating that centrality does not significantly moderate the relationship between salience and PAR. This suggests that, while centrality and salience each independently contribute to

perceptions of algorithm responsiveness, the effect of salience on PAR does not depend on the level of centrality.

These findings highlight the distinct roles of centrality and salience in shaping perceptions of algorithmic responsiveness. While both centrality and salience are positively associated with perceptions of responsiveness, their effects are independent of one another. Moreover, the strong impact of human-algorithm interplay and the negative role of ethical concerns further illustrate how user engagement and ethical considerations shape user experiences of algorithmic responsiveness.

Table 9: Centrality x Salience Interaction Term Regression Results

	<b>Coefficient</b>	<b>SE</b>	<b>95% CI</b>		<b>p</b>
			<b>Lower Limit</b>	<b>Upper Limit</b>	
<i>Constant</i>	22.62	0.37	21.89	23.35	.00
<i>Centrality</i>	1.25	0.39	0.49	2.02	.02
<i>Salience</i>	0.71	0.39	-0.05	1.47	.00
<b>Centrality x Salience</b>	0.64	0.35	-0.06	1.33	.49
<i>Content Filtering</i>	0.38	0.46	-0.54	1.29	.36
<i>H/A Interplay</i>	1.95	0.44	1.08	2.82	.00
<i>Automated Decisions</i>	-0.40	0.45	-1.28	0.49	.00
<i>Ethical Concerns</i>	-1.36	0.39	-2.13	-0.59	.00

### Human-Algorithm Interplay Mediation Analysis

The mediation analysis aimed to determine whether awareness of human–algorithm interplay—the recognition that algorithms personalize content based on users' identities and behaviors—mediates the relationship between racial identity centrality and perceived responsiveness. This aligns with Taylor and Choi's (2022) findings, which showed that PAR served as a mediator for medium enjoyment, enhancing user engagement by fostering perceptions of the algorithm's responsiveness to identity.

I tested whether interplay mediated the relationship between centrality and both PAR (Table 10) and PAI (Table 11) using the bootstrapping method (n=5000). None of these mediation analyses were significant. For both PAI and PAR, the indirect effects of centrality through interplay were close to zero with confidence intervals that included zero indicating no significant mediation. This means that interplay does not significantly mediate the relationship between centrality and perceptions of either PAR or PAI. In both cases, the total effect matches the direct effect due to the lack of significant mediation. These findings indicate that while centrality plays a direct role in shaping perceptions of PAR and PAI, human-algorithm interplay does not act as a mediating factor in this context.

Table 10: Centrality and PAR Interplay Mediation Results

<b>Effect Type</b>	<b>Coefficient (Variable)</b>	<b>Value</b>	<b>95% CI</b>	
			<b>Lower Limit</b>	<b>Upper Limit</b>
<i>Indirect Effect</i>	Centrality * Interplay	4.11E-16	-1.96E-15	1.99E-15
<i>Direct Effect</i>	Centrality on PAR	1.09	0.32	1.88
<i>Total Effect</i>	Centrality on PAR	1.09	0.32	1.88

Table 11: Centrality and PAI Interplay Mediation Results



<i>Effect Type</i>	<i>Coefficient (Variable)</i>	<i>Value</i>	<i>95% CI</i>	
			<i>Lower Limit</i>	<i>Upper Limit</i>
<i>Indirect Effect</i>	Centrality * Interplay	-4.02E-16	-1.90E-15	1.93E-15
<i>Direct Effect</i>	Centrality on PAI	1.09	0.30	1.88
<i>Total Effect</i>	Centrality on PAI	1.09	0.30	1.88

## Discussion

This study aimed to understand how Black TikTok users perceive algorithmic responsiveness (PAR) and insensitivity (PAI) in relation to their racial identity, focusing on racial identity centrality and racial identity salience within the framework of identity co-production. My findings indicate that racial identity centrality—a stable and intrinsic aspect of self-concept—significantly predicts both PAR and PAI. Users with higher racial identity centrality perceive the algorithm as both more responsive and more insensitive, reflecting a duality in algorithmic interactions where content may align with certain identity aspects but also reinforce or neglect others. Together, these results underscore that users with a strong racial identity centrality are more likely to experience both responsiveness and insensitivity, reflecting an ambivalent experience where the algorithm simultaneously aligns with and overlooks certain aspects of their identity.

Interestingly, racial identity salience—or the situational prominence of racial identity—was a positive predictor of PAR but did not significantly affect PAI. This finding aligns with MMRI literature, where centrality represents a stable sense of identity importance, while salience is dynamic and context-dependent. Salience may heighten the perception of responsiveness (PAR) when users feel their identity is acknowledged by the algorithm, but it does not necessarily influence insensitivity (PAI), which is more

tied to identity misrecognition or exclusion. My results suggest that heightened salience can lead to a more positive perception of the algorithm, especially when users feel recognized in ways that affirm their identity. These findings are consistent with research on adolescent ethnic-racial identity salience, which links situational identity awareness to positive psychological outcomes (Douglass et al., 2015). In contexts where racial salience is heightened, Black users may feel that the algorithm is responding more positively, aligning with the algorithmic crystal and the algorithmized self-concepts (Lee et al., 2022; Bhandari & Bhimo, 2021), which describes how algorithms dynamically reflect multifaceted aspects of the self in ways that encourage self-recognition and self-similarity.

Another key finding is that racial identity salience does not vary based on centrality in predicting perceived responsiveness. The moderation analysis showed no significant interaction between centrality and salience, suggesting that while both contribute to perceived responsiveness, salience acts independently of centrality. This highlights the unique roles of centrality and salience within identity co-production; centrality provides a stable sense of racial identity, while salience reflects situational changes. Thus, even users with lower centrality may feel increased responsiveness (PAR) when salience is temporarily heightened, such as in situations where their racial identity becomes more relevant.

This study further clarifies the distinction between algorithmic awareness and perceived responsiveness, showing that awareness—particularly human-algorithm interplay, or the understanding that users' actions can influence algorithmic recommendations—correlates significantly with PAR. This aligns with Taylor and Choi's (2022) findings, which indicate that while awareness and responsiveness are related, they remain conceptually distinct.

However, there are notable differences in findings related to PAI. Taylor and Choi found small negative correlations between PAI and both awareness of content filtering and human-algorithm interplay, suggesting that as users become more aware of algorithmic personalization, their perceptions of insensitivity may decrease slightly. In contrast, my regression analysis showed that content filtering awareness and human-algorithm interplay were significant positive predictors of PAI. This indicates that, in my sample, greater awareness of how algorithms filter and personalize content is associated with an increased perception of insensitivity. This discrepancy could suggest that while awareness of algorithmic processes may sometimes reduce perceived insensitivity, in other contexts, such as my study with Black TikTok users, this awareness heightens perceptions of insensitivity, possibly due to a heightened sensitivity to algorithmic biases or misrepresentation of identity – the experience of being ‘filtered out’ by the algorithm.

By framing these results within the identity co-production framework, this study contributes to understanding how racial identity and algorithmic systems mutually shape perceptions in digital spaces. Unlike human-algorithm interaction frameworks that emphasize user agency, identity co-production highlights the dual role of the user and algorithm. Users with strong racial identity centrality tend to interpret algorithmic content as both responsive and insensitive, which reflects how algorithms can simultaneously validate certain aspects of identity while disregarding others. This finding aligns with identity strainer theory (Karizat et al., 2021), which suggests that algorithms “filter” identity expressions based on implicit biases. High centrality users may see themselves as more consistently represented in algorithmic curation, but they are also attuned to moments when the algorithm fails to align with their identity, resulting in perceptions of insensitivity.

These findings align with concepts of algorithmic privilege and algorithmic representational harm, suggesting that users whose identities do not align with the algorithm's normative preferences may experience underrepresentation or misinterpretation. Identity co-production frameworks highlight how structural biases in algorithms shape user experiences in both positive and negative ways. This study supports the idea that racial identity centrality can amplify both responsive and insensitive experiences within the FYP, indicating that identity measures such as racial salience and centrality provide valuable nuance to our understanding of perceived responsiveness of the FYP algorithm.

### **Contributions, Limitations, and Future Directions**

This research fills a critical gap in algorithm studies by centering the experiences of marginalized users, specifically Black TikTok users, in examining perceived algorithm responsiveness. While Taylor and Choi's (2022) work on perceived algorithm responsiveness (PAR) and insensitivity (PAI) provides foundational insights, they acknowledge that their study's lack of diversity limits its scope and generalizability (p. 10). My study addresses this limitation, bringing essential nuance to the concept of algorithm responsiveness by exploring it through the lens of identity co-production, particularly for those whose identities are often underrepresented or misinterpreted by algorithmic systems. By focusing on how racial identity centrality and salience influence perceptions of responsiveness and insensitivity on the FYP, this study extends Taylor and Choi's conceptualization of algorithmic responsiveness beyond general user experiences to include the distinct, context-dependent interactions of marginalized users. This approach not only enhances our understanding of algorithmic responsiveness but also underscores the importance of examining identity-driven algorithmic interactions to capture the varied and complex ways that algorithms impact diverse user communities.

While this study provides meaningful insights, several limitations must be considered. The cross-sectional design limits understanding of how perceptions of PAR and PAI may change over time as users' interactions with TikTok evolve. Longitudinal research could illuminate the ongoing dynamics of identity co-production, particularly as algorithmic structures and user behaviors continue to adapt. Additionally, expanding the sample to include other identity groups and platforms would allow for a broader exploration of how algorithmic bias and responsiveness are experienced across diverse contexts and identities.

## **Conclusion**

This study contributes to understanding how Black users experience TikTok's FYP, emphasizing the unique roles of racial identity centrality and salience in shaping perceptions of algorithmic responsiveness and insensitivity. By situating PAR and PAI within an identity co-production framework, this research highlights the dual influence of user identity and algorithmic structures, demonstrating how algorithms simultaneously validate and misinterpret identity. The findings underscore the importance of recognizing both user agency and structural biases, affirming that algorithmic curation is not neutral but actively participates in shaping digital identity experiences. Identity co-production remains a critical framework for exploring how algorithmic systems reflect and reshape user identities in the algorithmic age, capturing the nuanced interplay between user perceptions, identity, and the algorithms that mediate these experiences.

## WORKS CITED

- Bhandari, A., & Bimo, S. (2022). Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society*, 8(1).
- Crasta, D., Rogge, R. D., Maniaci, M. R., & Reis, H. T. (2021). Toward an optimized measure of perceived partner responsiveness: Development and validation of the perceived responsiveness and insensitivity scale. *Psychological Assessment*, 33(4), 338–355.  
<https://doi.org/10.1037/pas0000986>
- Choi, M., & Toma, C. L. (2022). An Experiment on the Effects of Self-disclosure on Perceived Partner Responsiveness and Intimacy in Zero-Acquaintance Relationships. *Communication Studies*, 73(3), 297–313. <https://doi.org/10.1080/10510974.2022.2084429>
- DeVito, M. A. (2021). Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-38.
- Douglass, S., & Umana-Taylor, A. (2015). A brief form of the ethnic identity scale: Development and empirical validation. *Identity: An International Journal of Theory and Research*, 15(1), 48.
- Douglass, S., Wang, Y., & Yip, T. (2016). The Everyday Implications of Ethnic-Racial Identity Processes: Exploring Variability in Ethnic-Racial Identity Salience Across Situations. *Journal of Youth and Adolescence*, 45(7), 1396-1411.
- Eslami M, Rickman A, Vaccaro K, et al. (2015). "I always assumed that I wasn't really that close to [her]": reasoning about invisible algorithms in news feeds. In: CHI '15 proceedings of the 33rd annual ACM conference on human factors in computing systems, 153–162. New York: ACM.

- Gambino, A., Fox, J., & Ratan, R. (2020). Building a stronger CASA: Extending the computers are social actors paradigm. *Human-Machine Communication*, 1, 71–86. <https://doi.org/10.30658/hmc.1.5>
- Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic folk theories and identity: How TikTok users co-produce knowledge of identity and engage in algorithmic resistance. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 305, 1–44. <https://doi.org/10.1145/3476046>
- Lee, A. Y., Mieczkowski, H., Ellison, N. B., & Hancock, J. T. (2022). The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok. *Proceedings of the ACM on Human-computer Interaction*, 6(CSCW2), 1-22.
- Noble, S. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.
- Papacharissi, Z. (2011). A networked self. *A networked self: Identity, community, and culture on social network sites*, 304-318.
- Reis, H. T., & Shaver, P. (1988). Intimacy as an interpersonal process. In S. W. Duck (Ed.), *Handbook of personal relationships* (pp. 367–389). Wiley.
- Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A., & Chavous, T. M. (1998). Multidimensional model of racial identity: A reconceptualization of African American racial identity. *Personality and social psychology review*, 2(1), 18-39.

Shin, D. (2020). How do users interact with algorithm recommender systems? The interaction of users, algorithms, and performance. *Computers in Human Behavior*, 109, 106344.

<https://doi.org/10.1016/j.chb.2020.106344>

Simpson, E., & Semaan, B. (2020). For you, or for “you”? Everyday LGBTQ+ encounters with TikTok. *Proceedings of the ACM on Human-computer Interaction*, 4(CSCW3), 1–34. <https://doi.org/10.1145/3432951>

Taylor, S. H., & Choi, M. (2022). An initial conceptualization of algorithm responsiveness: Comparing perceptions of algorithms across social media platforms. *Social Media + Society*, 8(4).

<https://doi-org.proxy.library.upenn.edu/10.1177/20563051221144322>.

Volkova, K., & Kaun, A. (2019). Algorithmic Resistance: Social Media and Cultural Practices.

Zarouali, B., Boerman, S. C., & de Vreese, C. H. (2021). Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale).

*Telematics and Informatics*, 62, 101607.



# CHAPTER 4 : REFRAMING ALGORITHMIC RESISTANCE: EXPLORING ALGORITHMIC DISSONANCE IN BLACK USER ENGAGEMENT WITH TIKTOK'S FYP THROUGH MIXED METHODS ANALYSIS

## **Abstract**

This paper investigates the experiences of Black TikTok users to advance a framework of algorithmic dissonance, a concept describing the tension users experience when engaging with algorithms that reflect systemic racial biases while shaping their digital identities. Through a mixed-methods approach combining quantitative survey data and Critical Techno-cultural Discourse Analysis (CTDA), this study centers the often-overlooked voices of marginalized users, arguing that these experiences are central to understanding algorithmic influence. Moving beyond the traditional structure vs. agency debate, this paper proposes a frictive model of user-algorithm interaction, highlighting the dynamic interplay of self-concept and systemic bias within algorithmic spaces. By situating algorithmic dissonance as both a state of experience and a form of resistance, the study contributes to critical algorithm studies, algorithmic identity, and racial formation theory, calling for a more inclusive framework that positions marginalized voices at the core of algorithmic analysis.

## Introduction

In this paper, I introduce the concept of 'algorithmic dissonance,' which describes the tension users experience when engaging with personalization algorithms they simultaneously enjoy and view as potentially harmful or problematic. To develop this concept, I triangulate findings from two distinct but complementary studies. The first uses quantitative survey data to capture broad trends in Black users' perceptions of TikTok's For You Page (FYP) algorithm, while the second employs qualitative Critical Techno-cultural Discourse Analysis (CTDA) to explore how users articulate their experiences and perceptions in their own words through TikTok videos and comments. Though methodologically different, both studies are grounded in theories of algorithmic identity co-production and critical Black studies, providing a multidimensional view of how users navigate the platform. The quantitative data identifies patterns in perceptions and behaviors, while the qualitative analysis provides rich, context-specific insights that bring forward users' lived experiences. Together, these approaches reveal the complexities of 'algorithmic dissonance'—the cognitive and emotional tension Black users feel when engaging with an algorithm they recognize as flawed yet embedded deeply in their social media experiences.

Algorithmic dissonance further captures the passive or conflicted engagement users may have, where they continue using the platform despite recognizing its problematic aspects. Unlike traditional views of algorithmic resistance, which emphasize overt defiance or avoidance, algorithmic dissonance reflects the nuanced, everyday negotiations users undergo as they weigh the algorithm's harms against its benefits. This perspective expands the discourse, acknowledging that resistance does not always involve rejection; instead, it can coexist with compliance and enjoyment. This duality underscores the

complexity of user-algorithm interactions and challenges the binary view of users as either resisting or conforming to algorithmic systems.

As social media scholars have noted, TikTok offers a novel user experience that reveals how traditional understandings of social media culture and interactions are often oversimplified (Bhandari & Bimo, 2022). I argue that our understanding of interactions with the FYP algorithm is similarly limited. As algorithms evolve into more advanced forms of artificial intelligence, shaping daily interactions, media and technology theorists need foundational concepts that support theorizing around the complex and often contradictory ways users engage with these systems. Algorithmic dissonance offers such a framework, providing a lens to understand not only moments of resistance or perceptions of bias but also the fluid and dynamic relationships users have with algorithmic environments.

Identity scholars have begun to examine this relationship through a co-productionist lens, exploring how users' beliefs and perceptions of the FYP algorithm influence their behavior. Algorithmic dissonance extends this inquiry, capturing user beliefs, perceptions, and varied modes of engagement. While digital or algorithmic resistance often centers the algorithm itself, algorithmic dissonance shifts the focus to users. This approach advances discussions on user agency versus systemic biases, recognizing that users can simultaneously critique and depend on these systems, reflecting broader societal tensions surrounding race, identity, and technology. As algorithms become increasingly sophisticated and integrated into everyday interactions, it is crucial to move beyond the binary of user agency versus systemic bias, instead exploring the multifaceted ways users engage with personalized technologies.

To advance critical digital scholarship, I argue that bridging the agency-vs-structure binary is essential. Some scholars highlight how systemic biases permeate society, with race conceptualized as a technology, while others emphasize how users express themselves through digital affordances, where interface design shapes mediated communication. The cultural shift exemplified by self-making practices with the FYP algorithm calls for a concurrent understanding of how systemic biases manifest in user experiences and influence modes of self-expression.

Moving to my analysis, I outline each study's methods and findings, followed by a synthesis of insights from both. I then integrate my findings with relevant literature on digital resistance from critical Black digital studies and algorithm studies, establishing a framework that differentiates between algorithmic dissonance and algorithmic resistance. Both concepts are essential to understanding the current social media landscape: dissonance describes a specific user experience, while resistance encompasses a broader array of engagement forms. Finally, I discuss the value of mixed-methods research in capturing the increasingly complex digital lives we lead, underscoring how diverse approaches enhance our understanding of algorithmic systems and their impact on user experience.

### **Literature Review: User Agency v. Systemic Biases**

This literature review aims to show that the agency vs. structure debate no longer fully serves critical digital scholarship in an increasingly complex algorithmic landscape. I argue that critical theories, such as Racial Formation Theory, offer valuable perspectives for examining the intersection of user agency and algorithmic power. The traditional rift between agency and structure becomes blurred in algorithmic contexts, as algorithms complicate the boundaries between broader structural influence and

the day-to-day experiences and behaviors of users. Building on this complexity, the review introduces a reframing of the agency vs. structure tension as friction—where individual identity and systemic power interact continuously, creating points of resistance and adaptation rather than a simple dichotomy. By connecting these bodies of work through Racial Formation Theory, we can build a more robust framework for understanding how structure and agency are deeply interwoven within algorithmically driven environments.

### ***Race and Critical Digital Studies***

Historically, studies of institutional racism and structural inequality have only minimally addressed the influence of technology beyond mass media in shaping or challenging racial ideologies. Jessie Daniels argues that Internet studies often ignore the role of race in shaping online infrastructure and design, with technology histories frequently overlooking racial implications. In *Cyber Racism*, Daniels illustrates how white supremacy has adapted to digital platforms while recognizing marginalized groups' simultaneous use of the Internet for grassroots organizing and anti-racist efforts. Daniels challenges simplistic views of technology as either wholly harmful or beneficial, aligning with Nakamura's (2002) call for deeper consideration of race within internet studies. Daniels also argues that individual prejudice online receives more attention than the racism embedded in technological design, underscoring the need for studies that address algorithmic bias and discriminatory infrastructure.

The encoded biases within technology infrastructure offer a lens for understanding racial formation in the digital space, especially when examined through platforms designed with particular

racial assumptions. Safiya Noble's *Algorithms of Oppression* exemplifies this approach, drawing on theories such as Racial Formation, Mills' Racial Contract and Lipsitz's emphasis on the systemic over the individual. Noble argues that, in the digital space, discriminatory design is masked by the appearance of neutrality. She writes, "In the context of the web, we see the absolving of workplace practices such as the low level of employment of African Americans in Silicon Valley and the products that stem from it, such as algorithms that organize information for the public, not as matters of domination that persist in these realms but as democratic and fair projects, many of which mask the racism at play" (p. 80).

For Noble, underrepresentation of Black tech producers in Silicon Valley, coupled with the commercialization of the web, creates insensitive and racist algorithmic outputs. Lipsitz (1998) describes how these biases work to obscure coordinated, systemic behaviors within supposedly neutral technology, while Noble's analysis further illustrates how search engine technology replicates cultural stereotypes, limiting how Black users are represented. Reflecting on structural bias, Noble notes how these platforms maintain racialized hierarchies, hidden beneath claims of neutrality. Her view of algorithmic infrastructure as a "racial project" underpins the concept of friction-filled interactions, where users experience pushback as they encounter platforms' limited, racially biased representations. For Noble, simply being "present" on the web does not equate to agency, as online spaces are shaped by the limitations imposed by underemployment of Black tech producers in the industry. She notes, "The web can serve as an alternative space for conceiving of and sharing empowered conceptions of Black people" (p. 165), but argues that neoliberal policies restrict online spaces, shifting from collective rights toward individual freedoms.

This contention between systemic constraints and individual agency reflects the inherent tension in digital environments, where users navigate self-representation within a racialized infrastructure. Although Noble views technology as limited by design, I contend that the agency of Black tech producers could be leveraged to reimagine technology, paving the way for more equitable digital spaces. This potential is underscored by Noble's experience as an end-user, searching for a gift online only to be met with offensive results—a discovery that ultimately motivated her research. Her analysis suggests that user interaction with biased design may be seen as a friction-filled experience where racial identity meets structural bias. Noble emphasizes that these experiences only become visible when users with marginalized identities interact with the system, illuminating design flaws and sparking discourse about technology's impact on race and representation.

Noble's perspective aligns with Racial Formation Theory's assertion that race is a taken-for-granted reality that shapes self-perception and social experience. She underscores that issues of agency are intrinsically linked to issues of structure, showing how user interactions with algorithms reveal hidden biases. While Lipsitz argues that focusing on the individual obscures systemic bias, Noble's examples demonstrate that these individual interactions can reveal structural discrimination. This frictional relationship highlights that agency and structure are not isolated; rather, they interact within the same digital environments. Noble's view, in line with Lipsitz's framework, suggests that critical digital scholarship should approach agency and structure as coexisting forces within the same landscape.

Departing from Noble's focus on structural design, Nakamura (2007) explores agency through critical visual culture analysis, examining how individuals produce digital self-representations amid structural bias. Nakamura emphasizes the significance of agency within digital spaces, especially as users actively shape their identities. She writes,

It is crucial to examine users in their embodied subject positions, modes of production, especially amateur and low-end do-it-yourself digital cultures, which are usually more readily adopted by newer and less traditionally skilled or trained users(p. 6).

Her work foregrounds how these low-cost tools for digital self-representation empower users, countering structural biases. Writing before the era of social media dominance, Nakamura's focus on individual agency was ahead of its time, predicting how digital self-representation would become increasingly significant in online spaces. Her approach highlights how agency is neither separate from nor entirely constrained by structure but exists in a constant state of negotiation.

This friction between design and user interaction is further examined by Florini (2019), who considers how marginalized users engage with technology in ways that deviate from the designer's intent. Building on human-centered design principles, Florini argues, "Given that affordances are imagined through the interaction of the designers' intent, the materiality and functionality of the technology, and the perceptions, beliefs, and expectations of the user, then changing one of these components can yield radically different outcomes" (p. 184). Her perspective emphasizes the active role



of users in shaping digital spaces, revealing a continuous negotiation between user agency and technological structure. Rather than reinforcing binary distinctions, Florini's work illustrates how frictional interactions allow users to repurpose affordances in ways that address their unique needs and identities.

By placing emphasis on users' ability to adapt and redefine technology, Florini provides a counterpoint to Noble's focus on structural constraint, demonstrating that marginalized users actively interact with and challenge digital affordances. These experiences reveal a space of friction, where user agency and structural limits are constantly in flux. This view contrasts with Lipsitz's argument that focusing on the individual risks obscuring systemic bias. Instead, Florini suggests that examining individual agency reveals the points at which users push against structure, transforming their digital environments.

The review suggests that user engagement within algorithmically mediated spaces is better understood as a state of friction, where agency and structure continuously intersect and shape one another. Through this lens, the friction between user identity and algorithmic bias highlights both the constraints and possibilities within digital environments. By viewing these spaces as sites of negotiation, where marginalized users actively assert agency even as they navigate structural biases, this approach offers a more nuanced perspective on identity formation, resilience, and adaptation in the face of systemic digital constraints. This research, by integrating critical algorithm studies with Black digital scholarship, challenges the traditional binary of agency vs. structure as an analytical framework for understanding human-algorithm interaction.

### ***Algorithmic Identity and Theories of Race and Identity Formation***

The classic structure vs. agency debate presents a challenge for understanding identity in algorithmic spaces, especially when balancing the influence of TikTok's For You Page (FYP) algorithm with the user's agency. This debate is particularly relevant in examining racial identity within the FYP, where theories of racial formation provide valuable insights that reframe this tension as a dynamic, layered process. Race, understood here as both a personal identity marker and a societal force, is key to examining how algorithmically mediated environments shape interactions and self-concepts. The FYP algorithm tailors content to user behaviors and preferences rather than a social network, creating a feedback loop that intertwines individual expression with structural influences, which fundamentally complicates and enriches the structure vs. agency debate.

Initially, 'algorithmic identity' focused on categorizing users for targeted advertisements (Cheney-Lippold, 2011). However, recent scholarship has shifted toward understanding identity co-production from the user's perspective, emphasizing the role of algorithms in shaping and reflecting identity through a continual process of engagement. Co-production here involves ongoing reflection, where users not only experience the influence of algorithmic outputs but also actively shape their digital identities in response to them. TikTok users, in particular, navigate this interaction by reflecting on how the algorithm presents them and adjusting their behaviors in turn. This shift marks a movement beyond the 'networked self,' which emphasizes identity construction through social networks (Papacharissi, 2011). Instead, the 'algorithmized self' framework highlights continuous, algorithm-mediated identity work, centering the ways individuals engage with and adapt to algorithmic perceptions of themselves.

This interaction amplifies the user's role in actively negotiating their identity, shaped both by the algorithm's interpretations and the users' responses to these outputs.

This process of dual engagement—where users are both shaped by and shape their algorithmic representations—echoes theories of racial formation that similarly describe identity as both an individual and a structural construct. Omi and Winant's (1994) framework, for instance, presents race as a dynamic force that influences identity on both personal and societal levels, helping illuminate how racial identities are shaped within algorithmically driven spaces like TikTok. For Omi and Winant, race operates as a “sociohistorical process” encompassing both individual experience and systemic influences:

In contemporary popular discourse, it is often claimed that 'there is only one race—the human race.' Yet even as it is dismissed, race is also taken for granted. Though at times denied ethical legitimacy and scientific recognition, racial identity continues to constitute a fundamental aspect of human identity. How one sees oneself and how one is seen by others are both profoundly, and often contradictorily, shaped by notions of race. (p. 246)

This dual perspective on race aligns with the dual engagement described by the 'algorithmized self,' positioning TikTok users as participants in a complex interaction where their self-perceptions are mediated by the algorithm's reflections and interpretations of their identity. For Black users, this interplay becomes even more complex, as their digital and offline identities are both mediated by

historical and algorithmic perceptions of race. Thus, they navigate not only personal identity work but also the societal forces embedded within algorithmic systems that continuously reinforce certain racial narratives.

Hall (1997) offers a way of understanding the mechanism through which this sociohistorical process takes place. He argues that race operates more like a language than it does a predetermined biological marker. Hall (1997) posits that even though people do indeed have physical differences, race is a 'floating signifier' because its meaning is constantly evolving, not essential, and never fixed. He states,

[s]ignifiers refer to the systems and concepts of the classification of a culture, to its practices for making meaning. And those things gain their meaning, not because of what they contain in their essence, but in the shifting relations of difference, which they establish with other concepts and ideas in a signifying field (p. 4).

In digital spaces, this concept of race as a 'floating signifier' parallels how algorithms assign meaning to racial categories. The term 'signifier' also applies in human-computer interaction, where it refers to cues that communicate actions users can take. For instance, a magnifying glass icon signals search functionality, but there is nothing inherently universal about this symbol—its meaning is a social construct. Similarly, in algorithmic environments, algorithms interpret user data and behaviors, assigning meaning to racial identity in ways that reflect broader social and historical structures. This framework suggests that as users interact with TikTok's FYP algorithm, they engage in an ongoing negotiation where their racial identity is continuously interpreted and reinterpreted. The algorithm's influence prescribes and limits the ways in which racial identity is represented, leading to a co-

production of identity that is shaped both by personal expression and by algorithmic biases. This negotiation, particularly for marginalized users, reveals that identity in algorithmic environments is far more than a binary between agency and structure—it is a dynamic process in which personal and systemic dimensions of identity are constantly in dialogue.

Expanding on this view, the concept of the ‘algorithmic crystal’ refers to the way algorithms interact with user identity by reflecting self-concepts that are multifaceted, dynamic, and refracted through the identities of others encountered on the platform (Lee et al.). Within this framework, algorithms act like crystals that split user identity into multiple facets, returning back different versions that are shaped by both individual expression and algorithmic mediation. The algorithmic crystal also facilitates connections with similar others, orienting users to encounter facets of themselves through algorithmically curated content and groups. However, this theory often presumes a level of algorithmic neutrality, presenting the algorithm as a mirror or lens without considering how underlying structural biases may influence which aspects of identity are ‘refracted’ or amplified. This lack of critique overlooks how user identity is filtered through algorithmic systems embedded in a racialized hierarchy.

The Racial Contract, as Charles Mills (1997) outlines, provides a critical counterpoint by emphasizing that digital environments are not neutral spaces; rather, they are deeply embedded within the same racial hierarchies that shape offline society. Mills argues that race operates within a “partitioned social ontology” where whiteness is privileged, positioning non-white individuals as secondary or marginalized. This racial order functions as an “unspoken agreement” that upholds white supremacy by embedding racial hierarchies within social, moral, and political structures in ways that are

often unseen or unacknowledged. For Mills, individual agency within the Racial Contract is inherently constrained; though users can express themselves within the digital sphere, their interactions are filtered through structures that reinforce whiteness and render Black and non-white users as “sub-persons.” This racial contract is maintained through societal institutions, and digital algorithms are no exception; they reproduce social hierarchies, privileging dominant identities and limiting expressions of marginalized identities.

The critique here is that while the algorithmic crystal theory offers a compelling view of algorithm-mediated identity as multifaceted and fluid, it overlooks how algorithms operate within these racialized structures, treating them as neutral reflectors of identity. This assumption fails to address how structural biases within algorithms shape which facets of identity are reflected back to the user, potentially privileging certain identities while filtering out or distorting others. By assuming neutrality, the algorithmic crystal misses the role of the algorithm as an active participant in the racial contract, aligning more with Mills’ critique that such systems do not merely reflect identity but also maintain and reinforce racial hierarchies.

In algorithmic spaces, then, the refraction of user identity is not merely a product of individual interactions but is shaped by an algorithmic structure embedded in racial logic. TikTok’s FYP, for example, can be seen not as a neutral crystal reflecting all identities equally but as a site where white-centric norms are implicitly encoded, constraining the agency of marginalized users and influencing how they perceive themselves and others. The algorithmic crystal, when viewed through Mills’ framework, is

more accurately understood as a prism that privileges certain identities over others, aligning with and perpetuating the racial contract in digital form.

Expanding this critique, the Identity Strainer Theory offers a nuanced perspective on algorithmic bias by examining how individual user beliefs and perceptions can reveal broader, systemic biases within digital spaces like TikTok's FYP. While George Lipsitz (1998) warns against overemphasizing individual experiences at the expense of understanding systemic racial dynamics, Identity Strainer Theory challenges this perspective by illustrating how studying individual beliefs can contribute to a more comprehensive understanding of structural inequalities embedded in algorithms through identity co-production. Lipsitz argues that focusing on personal interactions alone can obscure the "sum total" of coordinated, systemic behavior that sustains racial hierarchies. However, folk theories like the Identity Strainer Theory harness individual user experiences to construct a broader theoretical framework, suggesting that individual perceptions of algorithmic bias provide key insights into the systemic mechanisms perpetuating racial hierarchies.

Identity Strainer Theory thus repositions individual experience as a window into systemic dynamics. This theory articulates how the FYP algorithm is perceived by users to 'filter out' marginalized identities based on race, body size, class, and other social markers. Users who fall outside the algorithm's biased framework experience 'algorithmic representational harm'—the negative effects of lacking algorithmic privilege, or the advantage held by those whose identities align with the algorithm's biases. By framing the suppression of marginalized identities as a form of algorithmic oppression,

Identity Strainer Theory illuminates how users actively interpret and resist the fyp's racialized filtering, making systemic bias more visible through individual accounts.

The Identity Strainer Theory extend Lipsitz's concerns by treating personal experiences not as isolated acts but as pieces of a collective narrative that illustrates systemic biases. This approach allows scholars to connect individual-level perceptions to a larger pattern of algorithmic discrimination, thereby acknowledging the structural while grounded in the experiential. By drawing attention to perceived algorithmic filtering, Identity Strainer Theory complicates the notion of neutrality and challenges the invisibility of systemic bias in algorithmic processes, building a bridge between individual experiences and broader social dynamics. This intersection suggests that, rather than obscuring systemic racial dynamics, individual experiences and beliefs about algorithmic bias can provide valuable insight into the structural forces shaping digital spaces, ultimately reinforcing the necessity of a dual approach that considers both personal and systemic dimensions of race within algorithmic identity formation.

In summary, examining algorithmic identity co-production through the lens of racial formation theories offers a robust framework for understanding how race operates within algorithmically mediated environments like TikTok's For You Page. Theories such as Omi and Winant's racial formation model, Lipsitz's critique of individual experience, and Mills's concept of the Racial Contract illustrate how race is both a personal and systemic construct. These frameworks reveal that the FYP algorithm is not a neutral tool but rather one embedded with racialized logics, where user interactions are shaped by historical and social forces that privilege certain identities while marginalizing others.



The Identity Strainer Theory further enriches this perspective by using individual user experiences to uncover broader systemic biases within algorithmic spaces. In doing so, it challenges Lipsitz's caution against focusing on personal interactions, demonstrating how individual beliefs and folk theories provide valuable insights into the structural biases woven into algorithms. By conceptualizing the FYP's filtering practices as a reflection of social inequalities, Identity Strainer Theory bridges the experiential and the systemic, arguing that user perceptions can illuminate the ways in which algorithms reinforce racial hierarchies.

This synthesis of algorithmic identity co-production with racial formation theories underscores the importance of analyzing both personal and systemic dimensions of race in digital spaces. It highlights the complex interplay between structure and agency in shaping identity within algorithmic environments, revealing that personal experiences are not merely isolated acts but part of a larger narrative that makes visible the systemic biases embedded in digital structures. Together, these insights contribute to a critical understanding of race and identity in algorithmic contexts, advancing a nuanced approach to studying how algorithms shape and reflect the sociohistorical realities of race.

### **Analysis: Finding Algorithmic Dissonance**

This section provides a brief overview of the methods of two studies designed to converge the data. I then detail how the data were triangulated to identify the concept of algorithmic dissonance.

#### ***Study One***

The first study investigates Black TikTok users' folk theories about the FYP algorithm, focusing on how they perceive the algorithm's interaction with their identities and how these perceptions shape

their behaviors and experiences on the platform. I conducted a Critical Techno-cultural Discourse Analysis (CTDA) of 139 TikTok videos and comments gathered through the search term "Black FYP algorithm." CTDA, developed by André Brock (2016), is an analytical approach tailored to studying internet phenomena, digital artifacts, and online culture. It combines analysis of the technological artifact with user discourse, framed by cultural theory, to examine the semiotic and material relationships between the form, function, belief, and meaning embedded in information and communication technologies. This method enables the integration of folk theories identified in the data into broader critical cultural research, highlighting algorithms as cultural artifacts that may reflect and reproduce racial logics.

Given the ever evolving and opaque nature of social media algorithms, developing theories on users' interpretations of algorithmic experiences is challenging, as user theories are frequently rendered obsolete without warning. DeVito (2021) addresses this issue by proposing the "process of folk theorization" framework, explaining how users continuously construct folk theories in response to the dynamic media environment—a particularly fitting approach for the personalized TikTok FYP algorithm. DeVito suggests that by examining the specific context of self-presentation, an enduring social process influenced by algorithmic mechanisms, we can assess the broader utility of folk theorization as a tool for understanding algorithms (DeVito, 2021, p. 2). This framework can enhance algorithmic literacy beyond a specific interaction and context, fostering a deeper, more adaptable comprehension of algorithms.

CTDA provides a robust framework for analyzing Black users' processes of folk theorization, situating their discourse within the critical perspectives of digital studies and Racial Formation Theory. This approach allows my analysis to contribute to theory-building in algorithmic identity, moving beyond

static user theories to provide insight into the underlying dynamics that shape behavior in today's media landscape. By examining how Black users articulate the influence of racial identity on their FYP experiences and behaviors, this study emphasizes their unique position at the intersection of internal identity co-production and algorithmic power, illuminating the complex interplay between user agency and algorithmic bias.

### ***Study Two***

In Study 2, I conducted a cross-sectional survey of 445 Black TikTok users to explore the relationships between perceived algorithm responsiveness, user algorithmic awareness, racial salience, and racial centrality. This study investigates how Black users perceive their racial identity as it relates to their interactions with the FYP algorithm and examines how racial identity measures influence perceptions of the algorithm's responsiveness to users' self-concepts. Racial centrality—the extent to which Black identity is core to a user's self-concept—was evaluated using a component of the Multidimensional Model of Racial Identity (MMRI) developed by Sellers et al. (1998).

The concept of perceived algorithm responsiveness draws from interpersonal communication theory, particularly the work of Taylor and Choi (2022), who adapt the traditional concept of responsiveness. In this context, responsiveness is understood as the extent to which “individuals come to believe that relationship partners both attend to and react supportively to central, core defining features of the self” (Reis et al., 2004, p. 203). Taylor and Choi's framework uses this concept to examine how algorithms might engage with users' self-concepts, transforming items like "my partner understands me" into "the algorithm understands me." Their framework introduces two components: perceived algorithmic responsiveness (PAR) and perceived algorithmic insensitivity (PAI). Perceived

algorithmic responsiveness (PAR) reflects the degree to which users feel that social media algorithms comprehend, acknowledge, and support essential aspects of their identities. Users reporting higher levels of PAR experience algorithmic content as more aligned with their self-concept and reasons for engaging with the platform. For example, a user with high PAR might feel that "the TikTok algorithm understands me." In contrast, perceived algorithmic insensitivity (PAI) measures the extent to which users feel that an algorithm disregards, limits, or misconstrues their identity and platform objectives. Users experiencing high PAI may feel that "the TikTok algorithm dismisses my interests." Together, these dimensions—PAR and PAI—illustrate the cyclical nature of perceived algorithm responsiveness: positive perceptions can enhance self-understanding and deepen platform engagement, whereas negative perceptions may prompt users to alter their behavior to better shape algorithmic responses.

Algorithmic awareness is another important, yet distinct, concept from perceived algorithm responsiveness. Users cannot gauge algorithm responsiveness if they lack an understanding of the algorithm's role in shaping their experience. Zarouali et al. (2021) address this need by developing the Algorithmic Media Content Awareness (AMCA) scale, a standardized tool applicable across various disciplines. The AMCA scale identifies four dimensions of user awareness. Content filtering awareness refers to users' recognition that algorithms curate media content for specific individuals based on online data. Automated decision-making awareness reflects users' understanding that algorithms independently decide what content to present. Human-algorithm interplay awareness refers to the user's insight into how their own behaviors affect the content that algorithms select to display. Finally, ethical considerations awareness encompasses users' acknowledgment of the ethical issues surrounding algorithmically recommended content. By exploring these dimensions, the AMCA scale enables a

comprehensive assessment of users' algorithmic awareness, capturing their understanding of how algorithms affect their content experiences.

Together, these measures of algorithm responsiveness and algorithmic awareness offer a rich foundation for understanding the nuances of Black users' interactions with the TikTok algorithm. They highlight how users' self-concepts, racial identity, and awareness of algorithmic influence coalesce in the FYP context, providing insight into the complex dynamics of identity co-production in a digital space shaped by racial and algorithmic logics.

### ***Data Convergence: Introducing Algorithmic Dissonance***

By synthesizing the data from Study 1 and Study 2, I introduce algorithmic dissonance as a critical concept that captures the conflicting emotions, beliefs, and engagement strategies of users who navigate an algorithm that both reflects and distorts their identities. This concept highlights the paradoxes inherent in algorithmic personalization for marginalized users and the tension between perceived control and perceived bias in digital environments.

The discourse analysis revealed that users felt validated when the algorithm seemed to "know" them, accurately reflecting their identity and serving content that resonated deeply. On the other hand, there was palpable frustration when the algorithm exhibited racial bias that reinforced stereotypes or failed to capture the multifaceted nature of their identity, leaving users feeling misrepresented or pigeonholed. This dual perception of bias highlights the intricate relationship between user identity and algorithmic behavior, where the lines between personalization, popularity, and bias blur, resulting in

feelings of both empowerment and alienation. Black users expressed a wide range of perspectives on how the algorithm interacts with their identities and developed strategies based on those perspectives.

The survey data further illustrate this duality, showing that users with higher racial identity centrality perceive the algorithm as more responsive overall, suggesting a sense of acknowledgment or recognition by the algorithm. However, these same users also report higher perceptions of the algorithm's insensitivity, highlighting the complexity of their interaction with the platform. While the algorithm seems to 'see' them, it simultaneously fails to fully understand or represent the depth of their identity. This paradox—the concurrent perceptions of responsiveness and insensitivity—serves as the foundation for understanding **algorithmic dissonance**, where users experience both empowerment and alienation from the same algorithmic system.

While awareness does play a significant role in shaping how responsive users perceive the algorithm to be, it is not the main reason individuals with high racial identity centrality see the algorithm as more responsive. Notably, both mediation and moderation analyses showed that neither overall algorithm awareness nor specific awareness of human-algorithm interplay mediated or moderated this relationship. This suggests that the influence of racial identity centrality on perceived algorithm responsiveness is direct in this context and is not significantly impacted by users' awareness of how the algorithm functions or their understanding of human-algorithm interactions. In other words, users who view their racial identity as central to their self-concept do not necessarily perceive greater algorithmic responsiveness simply because they are more aware of how it works. Thus, algorithmic dissonance does not conceptually rely on awareness alone.

Users with high racial identity centrality engage with TikTok's algorithm in ways that go beyond simple awareness, as evidenced by the three types of engagement patterns identified in my CTDA—tactical, experimental, and critical. These engagement modes help explain the relationship between racial identity centrality and perceived algorithmic responsiveness. Experimental engagement involves users testing their own folk theories to reveal algorithmic biases. Tactical engagement, on the other hand, consists of deliberate efforts to influence or manipulate the algorithm, such as following strategies to 'fix' the algorithm or align their content preferences with their identity, like achieving the 'Black girl algorithm.' This form of engagement ties their algorithmic identity more closely to their self-concept and seeks more control over their FYP.

Critical engagement reflects a more passive but reflective approach, where users interact with the FYP with heightened awareness of the algorithm's influence on their experience. Rather than actively trying to change the algorithm, they critically analyze its behavior, recognizing that the racial bias they encounter is not just a technical issue, but a reflection of larger systemic racism embedded in digital structures. This engagement mode shows users maintaining a critical distance, allowing them to question the algorithm without necessarily internalizing its biases. By doing so, they share insights and strategies with others to navigate the platform while avoiding harmful effects, thus fostering a more informed but less confrontational approach. These engagement patterns—experimental, tactical, and critical—reveal how users' perceptions of algorithmic responsiveness are shaped not just by their awareness of how the algorithm works, but by how they actively or passively engage with it. This complex interplay of strategies and reflections helps explain the paradox of users feeling both

empowered and frustrated by the algorithm, contributing to the broader concept of algorithmic dissonance.

Finally, ethical concerns about the algorithm's potential harms showed a significant negative relationship with both perceived insensitivity and responsiveness. While the negative relationship with responsiveness was expected, the negative association with insensitivity was surprising. The data suggest that users with higher ethical concerns do not view the algorithm as an independent creator or amplifier of bias, but rather as a tool that exacerbates existing societal biases, particularly those rooted in racism and anti-Blackness. These users perceive the algorithm as reinforcing harmful social agendas rather than acting neutrally. Given their critical engagement and awareness of the algorithm's potential harms, they not only perceive the algorithm as less responsive overall, but also less insensitive—because it functions exactly as they expect. Their understanding of the algorithm's role in perpetuating societal biases leads them to see it as predictable in its insensitivity, aligning with their critical perspective on its operations. This type of engagement reflects a more subtle form of algorithmic dissonance where it becomes the user's default state for engaging with the FYP.

In conclusion, the concept of algorithmic dissonance encapsulates the complex and often paradoxical relationship between users, particularly those with high racial identity centrality, and algorithmic systems like TikTok's FYP. Through both survey data and discourse analysis, it becomes clear that users experience simultaneous feelings of empowerment and frustration as they engage with an algorithm that both acknowledges and distorts their identities. The three engagement modes—experimental, tactical, and critical—demonstrate the various ways users navigate and respond to the algorithm's perceived responsiveness and bias, shaping their experiences on the platform. While



algorithm awareness plays a role in shaping perceived responsiveness, it does not fully explain the strong relationship between racial identity centrality and users' perceptions of the algorithm. Instead, users' critical engagement with the algorithm, combined with their understanding of its role in reinforcing societal biases, underscores the deep complexity of this interaction. This tension, where users see the algorithm as both a reflection of and contributor to systemic biases, helps to define the nuanced and multifaceted experience of algorithmic dissonance.

### **Discussion: Reframing Algorithmic Resistance**

Situating algorithmic dissonance within critical internet studies deepens our understanding of resistance in algorithmic systems, especially through users' lived experiences. By bridging algorithmic identity co-production with critical internet literature, I argue that algorithmic dissonance captures users' perception of Benjamin's (2019) concept of the New Jim Code, which reveals how racial biases are subtly embedded within digital design. This research thus extends racial logic into the digital era, where discriminatory design is more overtly experienced, particularly within personalized social media content feeds.

Benjamin describes the New Jim Code as the “backdoor” embedding of racial biases in technology design, often hidden from users. Through concepts like engineered inequity, default discrimination, coded exposure, and technological benevolence, she illustrates designs that, intentionally or not, reinforce social hierarchies and replicate societal divisions—even when aiming to reduce bias. Algorithmic dissonance captures users' cognitive and emotional responses to these concealed biases, revealing their awareness of engaging with algorithms that often fail to align with social equity goals. In this way, algorithmic dissonance represents users' understanding of the New Jim

Code, highlighting the tension they feel when engaging with personalized systems that embody systemic biases.

Hall's concept of the "signifier" is particularly relevant here; he describes signifiers as symbols or markers that convey meaning, which shifts based on social and cultural contexts. In algorithmic environments, this idea takes on new meaning as users communicate with algorithms through specific actions and affordances—such as likes, shares, and search behaviors—that act as digital signifiers of identity. Understanding algorithmic dissonance thus requires recognizing how the construction of racial identity in digital spaces is influenced not only by users' self-expressions but also by how algorithms interpret these digital signifiers within the limits of their design. Hall's notion of shifting signifiers aligns with users' engagement with algorithms, as they negotiate their identities through these affordances. Just as social contexts influence the meaning of racial signifiers, the digital environment constrains and shapes how these signifiers are processed and represented by algorithms, which can simultaneously shape and limit users' racial identities within algorithmic spaces.

While algorithmic dissonance theoretically extends our understanding of evolving racial logics, in practice, it does not fit neatly into traditional notions of resistance. Benjamin and Safiya Noble describe resistance as the active steps taken by individuals and communities to confront algorithmic discrimination, including raising awareness, building alternative technologies, and advocating for policy change. Noble's work emphasizes how marginalized groups challenge algorithmic oppression by creating spaces that counter these biases. Thus, algorithmic resistance encompasses a spectrum of actions aimed at recognizing and addressing systemic inequalities in digital spaces.

However, algorithmic dissonance itself is not a form of overt resistance. Instead, it reflects a state of heightened awareness—a cognitive and emotional tension regarding bias that influences engagement but is not rooted in protest or policy reform. While dissonance involves awareness of systemic bias, it is distinct from resistance, which often implies active opposition to that bias. On TikTok’s FYP, for instance, studies of algorithmic resistance focus on users’ attempts to influence the algorithm to produce content that aligns with their identities through strategies like “strategic refinement” or “polishing” (Karizat et al., 2021; Lee et al., 2022). Though some engagement strategies align with resistance, not all forms of engagement seek to align the algorithm with personal identities. Experimental and critical engagements, for example, differ in their goals, while tactical engagement involves advanced strategies beyond the affordances directly available in the algorithm. Engagement may reflect algorithmic dissonance, but dissonance itself does not require specific engagement; it is an experience rather than an act of protest.

Resistance within algorithmic spaces can also take the form of a subtle, ongoing friction between user identity and algorithm, giving rise to what I term algorithmic dissonance. This dissonance represents the tension users experience as they navigate the dual realities of self-representation and systemic bias in digital environments. Algorithmic dissonance exemplifies the dual nature of race, which operates simultaneously on individual and systemic levels within algorithmic systems, where personal identity consistently encounters structural bias. The friction inherent in algorithmic dissonance reflects the broader tension between agency and structure discussed in racial formation theories and critical digital studies, reframing the ongoing negotiation of identity in digital spaces as an act of resistance.

This dynamic is reminiscent of W.E.B. Du Bois's concept of "double consciousness," where Black users are acutely aware of how the algorithm perceives and categorizes them and must constantly navigate these perceptions to express an authentic sense of self. Algorithmic dissonance thus represents a kind of digital double consciousness, where users grapple not only with their own self-perception but also with the algorithm's biased interpretations of their identity. This dual awareness compels users to continuously negotiate how the algorithm frames their racial identity, sometimes subtly pushing back by carefully curating their interactions to better align with their true self-concept.

Furthermore, Algorithmic dissonance operates as both a state of being and a form of resistance. By engaging with the FYP algorithm while experiencing algorithmic dissonance, users perform a quiet resistance that challenges the algorithm's reductive assumptions. This friction underscores the complexity and resilience of identity formation in algorithmically mediated spaces, as users navigate and subtly reshape biased digital environments to preserve their agency and authenticity within systemic racial logic.

### **Conclusion: The End of Edge Cases**

This paper centers the experiences of marginalized users to advocate for a shift in critical algorithm studies, arguing against viewing marginalized identities merely as edge cases or anomalies in algorithmic design. Instead, it positions these lived experiences as essential to understanding how algorithms both reinforce and challenge societal structures. This approach highlights the structural biases embedded in digital systems and, importantly, reveals the nuanced ways in which users navigate and resist these biases in daily interactions. By focusing on the often-invisible forms of resistance

displayed by marginalized users, this study advances an inclusive framework for algorithm analysis, one that treats these voices as central rather than peripheral to the algorithmic experience.

In challenging the conventional framing of the structure vs. agency tension, this paper argues that a binary framework inadequately captures the complexity of algorithmic systems, particularly for marginalized users. Instead of a simple push-and-pull between structure and agency, interactions between users and algorithms are more accurately characterized as frictive—a clash within a shared space. In the case of algorithmic dissonance, users do not merely resist algorithmic structures nor act with complete agency. Rather, they navigate a landscape where their self-representations meet systemic racial logics, creating a nuanced interplay that transcends binary frameworks. This frictive model reveals that user engagement with algorithms is dynamic and continuous, marked by negotiations that shape both the user's self-concept and the algorithm's representation of them.

Reframing structure vs. agency as ongoing friction allows critical digital studies to more effectively analyze how identity, power, and resistance unfold in algorithmically mediated spaces. This shift acknowledges the layered experiences of marginalized users, who assert agency even as they encounter systemic constraints. The resulting friction is not merely oppositional but also productive, creating space for new forms of identity, resilience, and resistance. By placing marginalized voices at the center, this approach captures the intricate ways users adapt, resist, and redefine their engagement with algorithms, pushing critical digital scholarship beyond binary understandings toward a richer view of agency in algorithmic systems.

In introducing algorithmic dissonance, this paper contributes both theoretically and methodologically to the study of algorithmic identity and racialized digital spaces. Algorithmic

dissonance describes the persistent tension users experience when engaging with personalization algorithms that reflect systemic biases. This concept exemplifies how digital spaces like TikTok's FYP serve as sites where individual identity intersects with structural racial logic, prompting users to reconcile their self-perception with the algorithm's often reductive categorizations. By positioning resistance not merely as active opposition but as an ongoing negotiation of bias in the pursuit of authenticity and agency, algorithmic dissonance broadens the scope of digital resistance. This framework advances critical digital studies and racial formation theories by showing how resistance is continuously embodied in the lives of those navigating racially biased algorithmic systems.

Methodologically, the mixed-methods approach employed here strengthens these contributions by combining quantitative survey data with qualitative Critical Techno-cultural Discourse Analysis (CTDA) to provide a robust understanding of algorithmic dissonance. The survey data identifies trends in Black users' perceptions of TikTok's FYP, while CTDA offers interpretive depth, capturing users' lived experiences as they encounter and respond to algorithmic biases. This methodological blend enables a comprehensive view of algorithmic dissonance, illuminating the interplay between structural and individual identity formation in algorithmic spaces and underscoring the importance of interdisciplinary methods to fully explore digital identity co-production.

Together, these contributions establish a foundation for future research in algorithmic identity, critical race studies, and digital resistance, emphasizing the need to place marginalized voices at the core of algorithm studies. By conceptualizing algorithmic dissonance as both a state of experience and a form of resistance, this paper calls for a more inclusive approach to understanding identity, agency, and

resistance in algorithmic contexts, positioning the experiences of historically marginalized users as foundational to the study of algorithms rather than peripheral.

## WORKS CITED

- Abidin, C. (2021). Mapping Internet Celebrity on TikTok: Exploring Attention Economies and Visibility Labours. *Cultural Science Journal*, 12(1), 77–103.  
<https://doi.org/10.5334/csci.140>
- Barta, K., & Andalibi, N. (2021). Constructing Authenticity on TikTok: Social Norms and Social Support on the "Fun" Platform. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1–29. <https://doi.org/10.1145/3479574>
- Benjamin, R. (2019). *Race after technology: Abolitionist tools for the new Jim code*. Polity.
- Bhandari, A., & Bimo, S. (2022). Why's Everyone on TikTok Now? The Algorithmized Self and the Future of Self-Making on Social Media. *Social Media + Society*, 8(1).
- Brock, A. (2016). Critical technocultural discourse analysis. *New Media & Society*, 20(3), 1012–1030. <https://doi.org/10.1177/1461444816677532>
- Cheney-Lippold, J. (2011). A new algorithmic identity: Soft biopolitics and the modulation of control. *Theory, culture & society*, 28(6), 164-181.
- Daniels, J. (2009). *Cyber racism: white supremacy online and the new attack on civil rights*. Lanham, Md., Rowman & Littlefield Publishers.
- DeVito, M. A. (2021). Adaptive folk theorization as a path to algorithmic literacy on changing platforms. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), 1-38.
- Florini, S. (2019). *Beyond hashtags: Racial politics and Black Digital Networks*. NYU Press.



- Hall, S. (1997). Race, the Floating Signifier: What More Is There to Say about “Race”? In S. Hall, P. Gilroy & R. Wilson (Eds.), *Selected writings on race and difference*, 359-373. Duke University Press.
- Karizat, N., Delmonaco, D., Eslami, M., & Andalibi, N. (2021). Algorithmic folk theories and identity: How TikTok users co-produce knowledge of identity and engage in algorithmic resistance. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW2), Article 305, 1–44. <https://doi.org/10.1145/3476046>
- Lee, A. Y., Mieczkowski, H., Ellison, N. B., & Hancock, J. T. (2022). The algorithmic crystal: Conceptualizing the self through algorithmic personalization on TikTok. *Proceedings of the ACM on Human-computer Interaction*, 6(CSCW2), 1-22.
- Lipsitz, G. (1998). *The Possessive Investment in Whiteness: How White People Profit from Identity Politics*. Temple University Press.
- Mills, C. W. (1997). *The racial contract*. Cornell University Press.
- Nakamura. (2002). *Cybertypes: race, ethnicity, and identity on the Internet*. Routledge. <https://doi.org/10.4324/9780203953365>
- Nakamura, L. (2007). *Digitizing race: Visual cultures of the Internet*. U of Minnesota Press.
- Noble, S. (2018) *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press.

- Omi, M., & Winant, H. (1994). *Racial formation in the United States: From the 1960s to the 1990s* (2nd ed.). Routledge.
- Papacharissi, Z. (2011). A networked self. *A networked self: Identity, community, and culture on social network sites*, 304-318.
- Reis, H. T., & Shaver, P. (1988). "Intimacy as an interpersonal process." In *Handbook of Personal Relationships: Theory, Research and Interventions*, edited by S. Duck, D. F. Hay, S. E. Hobfoll, W. Ickes, and B. M. Montgomery, pp. 367–389. John Wiley & Sons, Oxford, England.
- Sellers, R. M., Smith, M. A., Shelton, J. N., Rowley, S. A., & Chavous, T. M. (1998). Multidimensional model of racial identity: A reconceptualization of African American racial identity. *Personality and social psychology review*, 2(1), 18-39.
- Taylor, S. H., & Choi, M. (2022). An initial conceptualization of algorithm responsiveness: Comparing perceptions of algorithms across social media platforms. *Social Media + Society*, 8(4). <https://doi-org.proxy.library.upenn.edu/10.1177/20563051221144322>.
- Volkova, K., & Kaun, A. (2019). Algorithmic Resistance: Social Media and Cultural Practices.
- Zarouali, B., Boerman, S. C., & de Vreese, C. H. (2021). Is this recommended by an algorithm? The development and validation of the algorithmic media content awareness scale (AMCA-scale). *Telematics and Informatics*, 62, 101607.

## CHAPTER 5: CONCLUSION

This dissertation makes significant contributions to the growing body of research at the intersection of algorithms and user identity by advancing the concept of algorithmic identity co-production—a framework that examines how users and algorithms collaboratively shape digital identities. Central to this work is the introduction of the concept of algorithmic dissonance, which captures the conflicting emotions and beliefs Black users experience regarding the TikTok For You Page (FYP) algorithm's racial bias. Algorithmic dissonance highlights the tension between the perception that the algorithm is biased against Black users and the enjoyment derived from TikTok's personalized user experience. This dissonance plays a crucial role in shaping how users navigate and negotiate their identities on TikTok, influencing not only their content creation and interaction patterns but also their broader perceptions of how the algorithm interprets their racial identity.

By moving beyond the binary framework of agency versus structure and integrating relevant studies, I emphasize the concept of algorithmic identity co-production, where Black users actively engage with and respond to algorithmic outputs that are shaped by broader sociocultural contexts. This perspective reveals that identity formation in algorithmically mediated environments is not merely about user agency or algorithmic control, but a dynamic process where users negotiate and contest their identities within the constraints and possibilities set by algorithmic systems. When the algorithm fails to deliver the personalized experience, it is designed to provide, users' interactions reveal the co-production of identity, challenging both the user's self-concept and the algorithm's role in representing

it. This analysis contributes valuable insights to the emerging interdisciplinary body of research on algorithmic identity by focusing on the experiences of users whose perspectives have not been deeply explored.

Furthermore, I demonstrate how algorithmic dissonance shapes user behavior and interactions with the TikTok FYP algorithm through three types of engagement: tactical engagement, experimental engagement, and critical engagement. Tactical engagement involves users sophisticated, high-effort maneuvering within the algorithm's parameters to maximize visibility; experimental engagement includes users testing the FYP algorithm based on their folk theories to uncover potential racial bias; and critical engagement occurs when users create content that critiques the algorithm and exposes its biases. By unpacking these dynamics, I provide a nuanced understanding of how Black users perceive and interact with the FYP algorithm in ways that are shaped by their racial identities and how these perceptions influence their behaviors. These types of engagement offer a deeper understanding of the relationship between racial identity centrality and perceived responsiveness, suggesting that users' perceptions are shaped by more than just awareness—they are influenced by the lived experience of navigating a racially biased algorithmic landscape.

Methodologically, this dissertation demonstrates the value of triangulating qualitative and quantitative approaches to understand the multifaceted ways in which users engage with algorithms. Through a combination of survey data and Critical Techno-Cultural Discourse Analysis (CTDA), I examine both the broader patterns in how Black users perceive algorithmic responsiveness and the specific discourses that reveal deeper layers of meaning behind these perceptions. The mediation analysis, using

Total Awareness as a mediator and incorporating all four dimensions of awareness, adds quantitative rigor to the study, while the discourse analysis provides context and depth that enriches the interpretation of the survey findings. This mixed-methods approach allows for a critical examination of how racial identity centrality and algorithm awareness intersect to shape perceptions of the FYP algorithm.

While the quantitative analysis revealed strong positive associations between racial identity centrality and perceived algorithm responsiveness, the mediation analysis showed that neither overall algorithm awareness nor specific awareness of human-algorithm interplay mediated this relationship. This finding suggests that algorithmic dissonance does not conceptually rely on awareness alone. Although awareness is relevant to perceived responsiveness, it does not fully explain the high perceived algorithm responsiveness among users for whom racial identity is central. The three types of user engagement identified in my CTDA—tactical, experimental, and critical—begin to unpack the relationship between racial identity centrality and perceived responsiveness, offering a way to interpret these associations. However, the regression analysis did not include the variables needed to fully explain the underlying mechanisms driving this relationship, suggesting that other factors, perhaps related to the nuances of identity expression and resistance, may be at play.

By bridging critical algorithm studies with critical race and digital studies, this dissertation advances the concept of algorithmic identity co-production as a dynamic and evolving area of research within the ever-changing algorithmic landscape. It underscores the importance of considering both algorithmic structures and sociocultural contexts that shape user experiences, moving beyond binary

frameworks of agency versus structure to explore how identities are actively co-produced in algorithmically mediated environments. This work invites future research to further explore these complex intersections, especially as digital platforms and algorithms continue to evolve and impact how marginalized users interact, resist, and negotiate their identities in digital spaces.

## APPENDICIES

### APPENDIX A: SURVEY INSTRUMENTS

*The Prolific screener asked demographic questions and participant's TikTok usage, heavy usage qualified as once per week. (R) indicates a reverse coded item. All questions were 5-point Likert scale (1 = not at all, 5 = completely.*

#### Centrality Scale

*How important Black identity is to users' self-concept*

1. Overall, being Black has very little to do with how I feel about myself (R)
2. In general, being Black is an important part of my self-image
3. My destiny is tied to the destiny of other Black people
4. Being Black is unimportant to my sense of what kind of person I am (R)
5. I have a strong sense of belonging to Black people
6. I have a strong attachment to other Black people
7. Being Black is an important reflection of who I am
8. Being Black is not a major factor in my social relationships (R)

#### Awareness Measures

*Content filtering represents baseline awareness that media content is personalized via algorithms, whereas human–algorithm interplay represents a more sophisticated awareness about the use of “big data” to automatically personalize content curation.*

Please indicate to which extent you are aware of the following statements about the FYP

#### **Content Filtering**

*Being aware that algorithms are used to tailor media content to specific people based on online data*

1. Algorithms are used to recommend post to me on my FYP

2. Algorithms are used to prioritize posts on my FYP
3. Algorithms are used to tailor certain posts on my FYP
4. Algorithms are used to show someone else different posts than I get to see on my FYP

### **Human-Algorithm Interplay**

*Being aware that one's individual behavior influences which content algorithms decide to*

*Present*

1. The videos on my FYP depend on my behavior on TikTok
2. The videos the FYP algorithm recommends to me depend on my behavioral data
3. The videos the FYP algorithm recommends to me depend on the data that I make available online

### **Automated decision-making**

*Being aware that algorithms are used to make automated decisions in tailoring media content to people*

1. Algorithms are used to show me videos on my FYP based on automated decisions
2. Algorithms do not require human judgments in deciding which videos to show me on my FYP
3. Algorithms make automated decisions on what video I get to see on my FYP

### **Ethical Considerations**

*Being aware of the ethical concerns of algorithmic-recommended content*

1. It is not always clear why the FYP algorithm decides to show me videos
2. The videos that the FYP algorithm recommends to me on TikTok can be subjected to human biases such as prejudices and stereotypes
3. Algorithms use my personal data to recommend certain content on TikTok, and this has consequences for my online privacy



## Saliency

*The degree to which users consider race on the FYP*

1. When I go on my FYP, I always take note of the racial make-up of the creators in my feed.
2. When I go on my FYP, I always look for content that is relevant to my racial identity.
3. I think that the algorithm is aware of my racial identity.
4. The FYP algorithm has shown bias based on my racial identity.

## Responsiveness Measures

*How the user perceives the algorithm responding to their self-concept*

I am interested in what you think about the TikTok 'for you page' (FYP) algorithm. Think about the posts, videos, and stories the algorithm curates for you, and then answer the following questions.

### **Perceived Responsiveness**

The FYP algorithm really listens to me

The FYP algorithm responds to what I am thinking and feeling

The FYP algorithm understands me

The FYP algorithm tries to see where I'm coming from

The FYP algorithm is attentive to my needs

The FYP algorithm takes my interests seriously

The FYP algorithm really gets who I am

### **Perceived Insensitivity (R)**

The FYP algorithm does NOT understand my feelings and concerns.

The FYP algorithm ignores who I am

The FYP algorithm dismisses my interests

The FYP algorithm seems to ignore the things that are most important to me

The FYP algorithm does NOT really understand my wants and needs

The FYP algorithm does NOT really take my personal interests seriously

The FYP algorithm often does NOT listen to my needs

## APPENDIX B: CENTRALITY SCALE FACTOR LOADINGS

<i>Item</i>	<i>Factor 1</i>	<i>Factor 2</i>
<i>1. Overall, being Black has very little to do with how I feel about myself (R)</i>	0.57	0.44
<i>2. In general, being Black is an important part of my self-image</i>	0.69	-0.11
<i>3. My destiny is tied to the destiny of other Black people</i>	0.45	-0.12
<i>4. Being Black is unimportant to my sense of what kind of person I am (R)</i>	0.65	0.35
<i>5. I have a strong sense of belonging to Black people</i>	0.76	-0.24
<i>6. I have a strong attachment to other Black people</i>	0.80	-0.31
<i>7. Being Black is an important reflection of who I am</i>	0.79	-0.15

8. *Being Black is not a major factor in my social relationships (R)*

0.49

0.39

### APPENDIX C: BOX PLOTS PER VARIABLE AND Z-SCORE OUTLIER ANALYSIS

FIGURE C1: CENTRALITY BOX PLOT

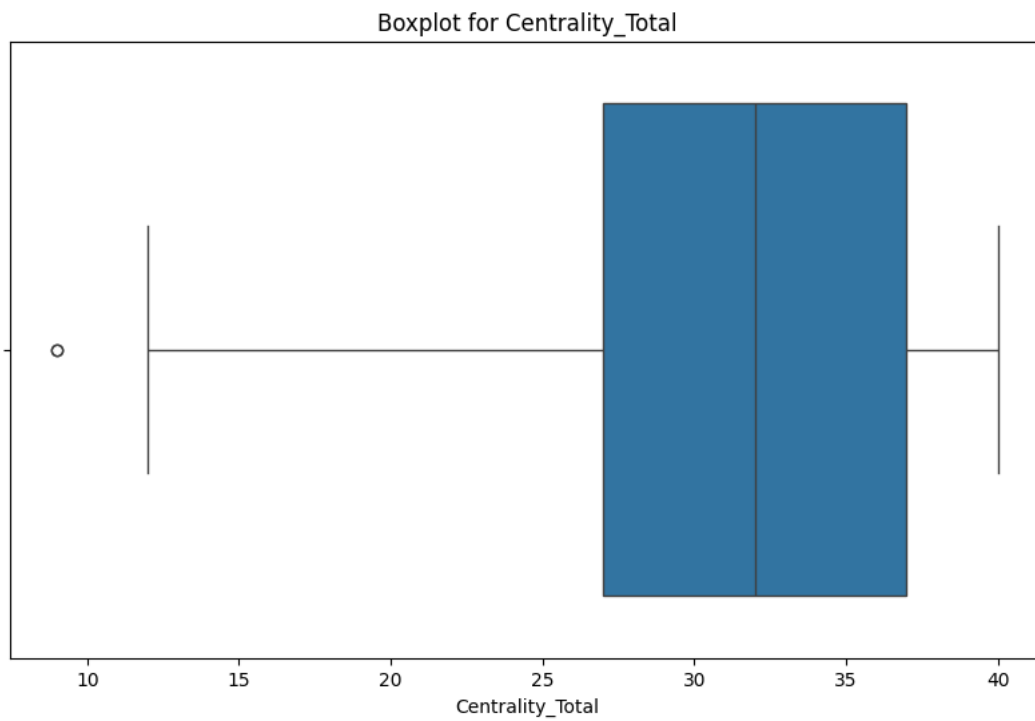


FIGURE C2: H/A INTERPLAY BOX PLOT

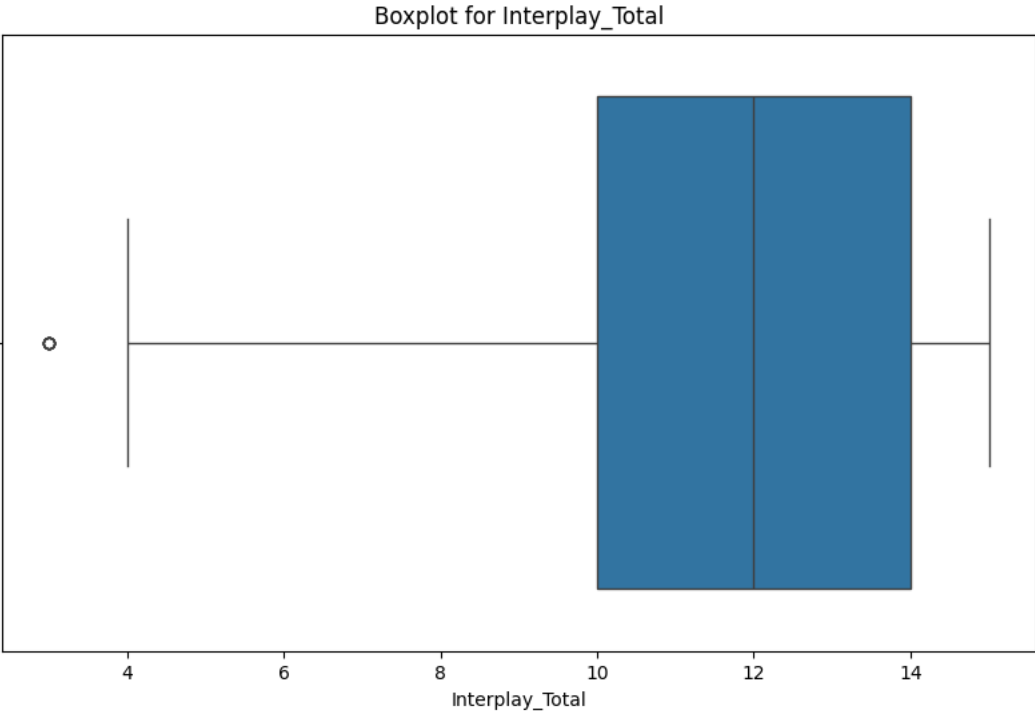


FIGURE C3: SALIENCE BOX PLOT

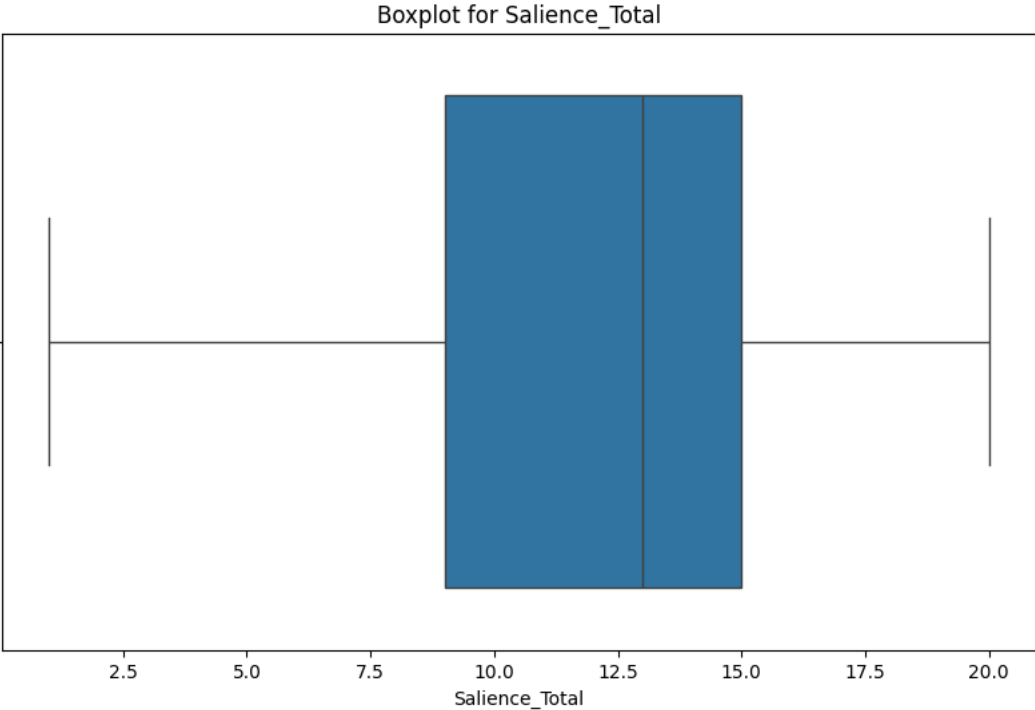


FIGURE C4: CONTENTFILTERING BOX PLOT

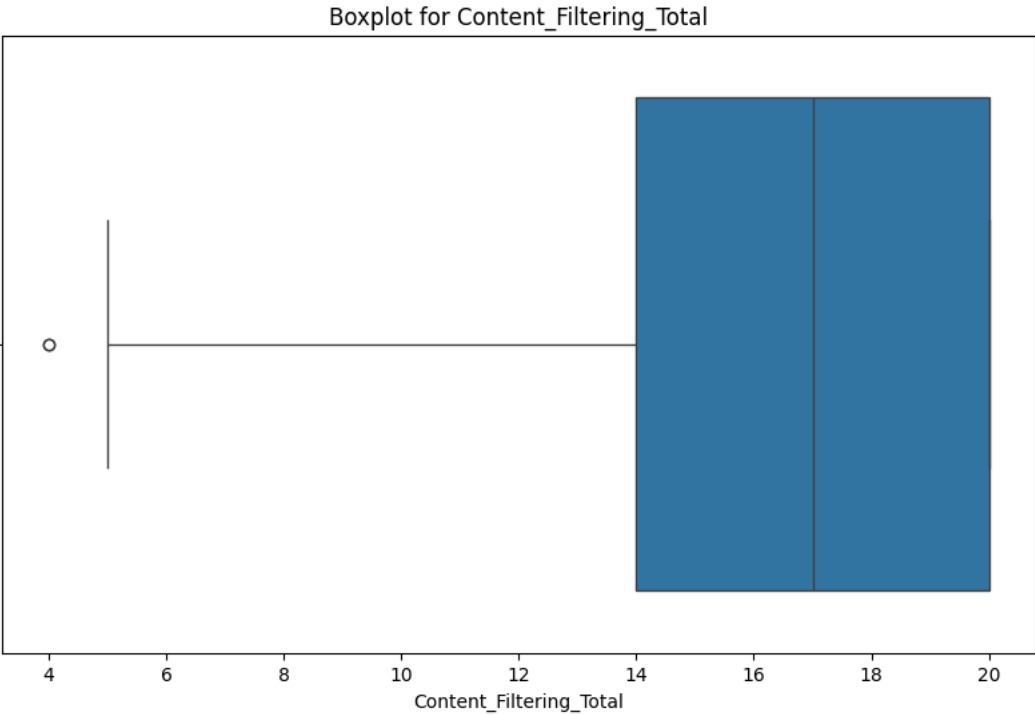


FIGURE C5: AUTOMATED DECISIONS BOX PLOT

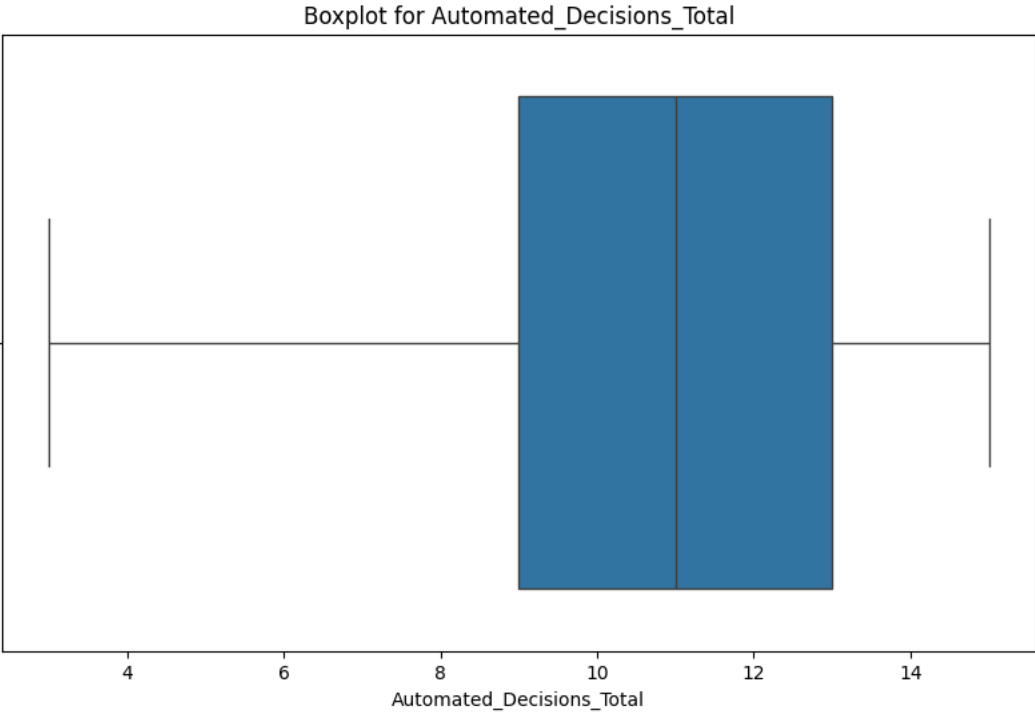


FIGURE C6: ETHICAL CONCERNS BOX PLOT

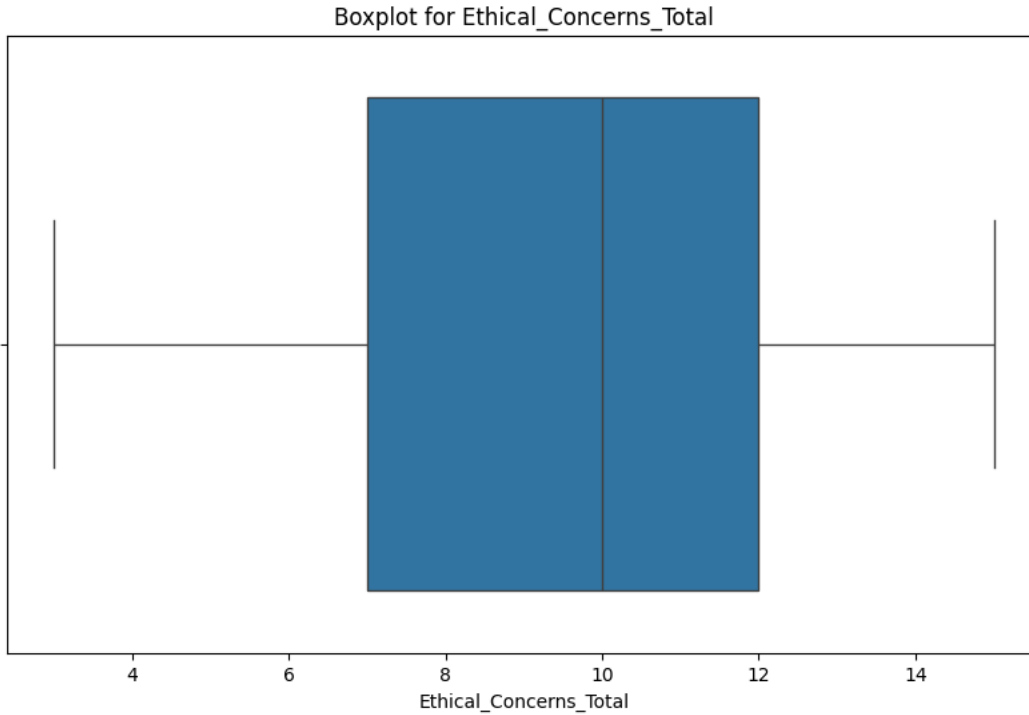


TABLE C1: Z-SCORE OUTLIER ANALYSIS

<i>Variable (standardized)</i>	<i># of Outliers</i>
<i>Centrality_Total_z</i>	2
<i>Saliency_Total_z</i>	0
<i>Content_Filtering_Total_z</i>	6
<i>Interplay_Total_z</i>	0
<i>Automated_Decisions_Total_z</i>	0
<i>Ethical_Concerns_Total_z</i>	0

TABLE C2: Z-SCORE OUTLIER ANALYSIS T/F THRESHHOLD

If a specific case has a Z-score beyond the threshold (typically |3|), it would be flagged as True for that variable, indicating an outlier. In your output, "False" means that for each case and variable combination, none exceeded the outlier threshold, and thus no extreme values were detected for that variable.



<b>Centrality</b>	<b>Saliency</b>	<b>Content Filtering</b>	<b>H/A Interplay</b>	<b>Automated Decisions</b>	<b>Ethical Concerns</b>
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
FALSE	FALSE	FALSE	FALSE	FALSE	FALSE