From Regret to Subscription: The Consequences of AI Classification Failures on Streaming Platforms

Ana Rita Gonçalves
Nova Information Management School
Diego Costa Pinto
NOVA IMS Lisbon
Hector González
ESCP Business School

From Regret to Subscription: The Consequences of AI Classification Failures on Streaming Platforms

Abstract

Smart services are increasingly using artificial intelligence (AI) to tailor recommendations and content. Drawing from AI classification experience and identity-based motivation, this research explores the impact of AI classification failures on consumers' self-identity connection and regret,and how a creative identity expression shapes this impact. Across three studies, this research shows that AI classification failures can reduce consumers' self-identity connection and have downstream effects on consumer behavior. Consistent with our identity-based account, AI's detrimental effects are amplified when consumers are motivated to use AI to express their identity and are mitigated when self-expression motives are not salient. Given the potential impact of AI classification failures to be more detrimental to customers' self-identity in this environment, our findings have significant theoretical and managerial implications for the developing field of smart service failures.

Keywords: AI Classification; self-identity; service failure

Track: Consumer Behavior

1. Introduction

Artificial Intelligence (AI) has reshaped many industries, as AI-powered recommendations have become a must-have to increase the revenues of digital services (Huh et al., 2023). However, despite the notable advantages of AI for smart services, it does not always answer consumers' expectations, leading to AI-enabled service failures (Grégoire & Mattila, 2021). Moreover, even though the self is integral to consumers' consumption behaviors (Gonzalez-Jimenez, 2017), the literature on service failure-recovery (SFR) does not offer much insight into AI failure's effects on consumers' sense of self.

Drawing on AI classification experience (Pantone et al., 2021) and identity-based motivation (Oyserman, 2009), we posit that AI recommendation failures can make consumers engage in self-defensive mechanisms, reducing consumers' self-identity connection and increasing negative outcomes such as regret, defensive neurophysiological cues, and adverse downstream behaviors for service providers. Consistent with our theorizing, we propose that when a creative identity expression is salient (vs. control), it shapes how smart service failures are processed. In particular, we posit that a creative mindset may further accentuate their regret after a smart service failure. Consequently, we propose that when consumers are motivated to use AI to express their identity, failures in AI classification can have downstream effects such as regret of using AI and reduced subscription behaviors. Across four studies, this research provides important theoretical and practical implications by bridging the AI classification experience (Puntoni et al., 2021) and identity-based motivation (Oyserman, 2009) literature by examining the mechanism of self-identity connection and by exploring its impact on consumers' self-expression processes related to AI classification experience failures. In addition, we extend the smart SFR literature (Choi et al., 2020) by exploring the consequence of AI failure for consumers' self in a timely yet scant research context, namely, online streaming. Finally, we provide guidelines to service providers.

2. Theoretical Background

A service failure refers to an interaction where a consumer experiences a loss due to a service provider's failure or when the provider fails to live up to customers' expectations, leading to negative behavioral outcomes (Patterson et al., 2006). However, there is limited understanding of smart SFR' effects on consumers' sense of self. In the context of our research, consumers will rely on an AI to make adequate content recommendations that align with their expectations. Recent tech advances in streaming have encouraged consumers to self-express through the creation of

assortment (e.g., Flecha-Ortíz et al., 2021) and to join in engaging activities when motivated by their need for self-expression (de Vries et al., 2017). Hence, self-expression has become key in AI-based streaming platforms for consumers to create such "self-expressive assortments" (Rifkin & Etkin, 2019). Consumer's self-identity is key to driving consumers' consumption choices (Gonzalez-Jimenez, 2017). Identity influences how we process information, react to certain objects, and participate in specific activities (Kirmani, 2009). However, if the AI provides a content selection that is not in line with expectations (i.e., fails on their classification), consumers will feel misunderstood (Adam et al., 2021). Consumers may feel that the AI has inaccurately assigned them to a wrong group, which can lead to identity conflict (Reed II et al., 2012). Arguably, the consumer is likely experiencing regret because they relied on the recommendation provided by the AI. Aligned with Puntoni et al. (2021), classification experiences may cause customers to feel misled if they believe AI incorrectly allocated them to a group or made biased predictions based on group assignments. This classification failure can generate regret in the consumer, which is associated with consumer self-identity (Davvetas & Diamantopoulos, 2017). Meanwhile, the literature suggests that beyond products, entertainment options such as movies are also considered self-relevant (Rozenkrants et al., 2017). AI's consumer categorizations can aid self-validation if identity motives are met, but failure scenarios may cause consumers to feel AI usage may not align with their self-identity. The human propensity for categorical thinking in person- and self-perception amplifies such assumptions (Turner and Reynolds 2011). Thus, we propose that self-identity connection is the underlying mechanism of smart service failures. A central tenet of this research is that consumers use movie choices to showcase their uniqueness and individuality and, consequently, self-express. This assertion derives from the main notion that a person's behavior is driven by a desire to reaffirm their self-image (Dunning, 2012). As such, a salient creative identity expression mindset may be a powerful driver of a consumer's perceptions and behavior because consumption can be identity-based, being both self-constructive and self-destructive (Oyserman, 2009). Mindsets refer to specific sets of activated cognitive procedures that result in particular ways of processing information (Buttner et al., 2013). A particular mindset can influence the products consumers are drawn to and the messages they find most persuasive (Murphy & Dweck, 2016). In our research context, we propose that the salience of an identity expression mindset may further accentuate their regret after a smart service failure. Indeed, research argues that identity expression (e.g., creativity) is integral to a person's sense of self (Czerwonka & Karwowski, 2018). When customers are under a self-expressive mindset, it amplifies the negative consequences of a failure to provide self-relevant content.

3. Overview of Studies

Three experimental pre-registered studies test our predictions. The first study reveals that AI classification failure can reduce consumers' self-identity connection, leading to regret and reduced subscription behaviors. The second study reveals that identity expression mindsets can exacerbate regret after AI classification failures. The third study uses neurophysiological evidence for objective emotional assessment.

Study 1: The Underlying Mechanism of Self-Identity Connection

Two hundred nine streaming users were recruited online (Prolific, 53.1% women, M_{age}=36.46; SD_{age}=13.742), and it employed a one-factor between-subject design.

Procedure, Stimuli, and Measures. Participants were randomly assigned to two experimental conditions. They were presented with an AI-powered tool that suggested three trailers based on their preferences. The success (vs. failure) condition showed trailers aligned (vs. contrary) with their choices. Then, the extent to which participants regret using the AI tool was measured using a 3-item scale adapted from (Tsiros & Mittal, 2000): "*I feel sorry for using AI tool", "I regret using the AI' tool", and "I shouldn't have used AI' tool.* The participants rated on a 9-point scale (1= Strongly disagree to 9= Strongly agree). As a form to understand what compartmental behavior of the participants after would be experiencing the AI tool, it was asked if they would subscribe to that AI tool. An answer of yes or no was provided. Self-Identity Connection was measured using a scale adapted from (Huang & Philp 2022): *"The AI tool content reflects part of me and who I am"; "I feel personally connected to the trailers presented by AI tool*". As for manipulation checks, the respondents were asked to indicate if it was a success or a failure (1=Success to 9=Failure).

Results. Manipulation Checks. The Independent Samples T-test results suggest that the failure manipulation worked as expected. Participants in the failure condition reported higher levels of failure ($M_{failure}$ = 5.89, $SD_{failure}$= 2.933) in comparison to participants in the success condition ($M_{success}$ = 3.57, $SD_{success}$ = 2.316; $t(207)$ =-6.358, $p < .001$).

Regret. Independent-Samples T-Test showed that participants in the failure condition presented higher regret values than participants in the success condition ($M_{failure}$=3.62 vs. $M_{success}$=2.33; $t (207)$ =-4.134; $p < 0.001$).

Subscription Behavior. A logistic regression was used to test if the participants in the success vs. failure conditions would subscribe more or less to these AI applications. The results show that

participants in the failure condition subscribed less than those in the success condition ($\chi^2(1)$ = 8.776, p= 0.003).

Mediation Effect of Self-Identity Connection on Regret. A mediation analysis was conducted using Hayes Process (model 4, Hayes, 2017; n=5000). The mediator was self-identity connection, the independent variable was failure vs. success, and the dependent variable was regret. The bootstrap analysis reveals the significant mediation effect of failure vs. success on regret (indirect effect (a × b) = 0.3917; 95% CI: 0.1486 to -0.6612) and significant direct effect on regret was not significant (direct effect [c] = 0.8936, p = .0028).

Mediation Effect of Self-Identity Connection on Subscription Behaviors. Mediation analysis was conducted using the same procedure as before, where the mediator was self-identity connection. The bootstrap analysis reveals the significant mediation effect of failure vs. success on subscription behavior (*indirect effect (a × b) = 0.5570; 95% CI: 0.2018 to 0.9710*) and no significant direct effect (*direct effect [c] = 0.4750, p = .1527*) revealing full mediation by self-identity connection.

Study 1 reveals that AI classification failure can decrease consumers' self-identity connection, leading to regret and decreased subscription behaviors. The study suggests that failure vs. success influences how self-identity connection impacts consumers' regret. Consumers exposed to failures had higher self-identity connection, leading to higher regret and decreased subscription behaviors, aligning with previous research on identity conflict.

Study 2: The moderating role of Creative Identity Expression

The study recruited 202 participants using Prolific. Participants from various streaming platforms were asked to indicate their preferred platform and complete a questionnaire. There were 202 participants (71.8% women, M_{age}=33.46; SD_{age}=12.39). The study used a 2x2 between-subjects design (identity expression salient vs. control, AI failure vs. success).

Procedure, Stimuli, and Measures. Participants were randomly assigned to a creative identity expression vs. control mindset condition, focusing on the importance of creativity in their lives. In a second phase, they were given a new AI tool on their favorite streaming platform, which would present three suggestions. In the success content condition, the suggestions aligned with users' preferences, while in the failure condition, the tool presented content different from what they typically watch to introduce them to new options. Afterward, regret was evaluated using the same scale as before. As for manipulation checks, the participants had to answer if they had to write about

being creative vs. focused and whether the suggestions from the AI tool were based on their preferences or not.

Results. Manipulation Checks. Results from two-way ANOVA provide support for the main effect of creative identity expression salient vs. control manipulation (F(1, 201) =215679.084, p < 0.001, ηp2 = 3170.048). Moreover, results show a main effect of the AI classification manipulation (F(1, 201) = 40.085, p < .001, ηp2 = 239.959) and no significant interaction effect between identity expression and failure relevance (F(1, 201) = 0.854, p = 0.356, ηp2 =0.002).

Regret. A two-way ANOVA using 2 (AI classification success vs. failure) X 2 (identity expression salient vs. control) on regret shows the predicted interaction effect on regret (*F*(1, 201)=4.095, *p* <.044). Specifically, when participants experienced a failure, it resulted in higher regret (*M*$_{failure}$=3.40, *SD* =2.086) than when they got a success (*M*$_{success}$ =2.19, *SD*=1.755) when a creative identity expression was salient. However, in the control condition, participants experiencing a failure or success yielded similar results of regret (*M*$_{failure}$ =2.95, *SD*=2.109; *M*$_{success}$ =2.930, *SD*=2.304). In addition, results show a significant main effect of AI classification (F(1,201)=4.426, p=.037) and no significant main effect of creative identity expression (F(1,201)=0.240, p=.625).

The study reveals that a creative identity-expression mindset, fostering creativity and not controlling, significantly impacts how smart service failures are processed. This mindset amplifies the negative consequences of classification failures, particularly when users' preferences are self-irrelevant. This mindset leads to greater regret when using AI tools for recommendation, as users are motivated to act in identity-congruent ways. This can result in incongruent identity recommendations, making consumers feel understood and potentially leading to ineffective service delivery.

Study 3: Facial Expression Recognition Software and Self-Identity

Participants were recruited from a large European university. The sample consisted of one hundred participants (69.9% women; Mage=24.69, SD=5.924). The study employed a single-factor AI classification failure vs. success between subject design.

Procedure, Stimuli and Measures. A study at a European University used facial expression recognition software to analyze participants' facial expressions. Participants were randomly assigned to two experimental conditions: one where they imagined they were on their favorite streaming platform, and the other where the AI tool suggested a trailer based on their preferences. The success condition showed a trailer aligned with their choices, while the failure condition showed a trailer contrary to their preferences. The study included romance, comedy, action, and animation movies,

selected to not be on the cinema by May 2023. With the use of the Facereader, it was possible to include facial expression analysis, emotion data and offers reliable automated facial action coding that was also used, as well as valance and arousal. Self-Identity Connection was measured using a scale adapted from (Huang & Philp, 2022) as before. A binominal question to measure behavioral intentions of like and dislike was also included. Finally, as for manipulation checks, the respondents were asked to indicate if the experience with the AI tool was a success or a failure (1=Success to 9=Failure).

Results. Manipulation Checks. The results from Independent Samples T-test table suggest that the failure manipulation worked as expected ($M_{failure}$ = 4.71, $SD_{failure}$= 2.602; $M_{success}$ = 2.84, $SD_{success}$ =1.821; t(99) = -4.474, p < .001).

Sad. Independent-Samples T-Test showed that participants in the failure condition presented higher sad emotions data than participants in the success condition ($M_{failure}$ = 0.12 vs. $M_{success}$ =0.07; *t (99)* =-2.423; *p* = 0.008).

Arousal and Valance. Independent-Samples T-Test showed that participants in the failure condition presented lower levels of arousal and valance than participants in the success condition ($M_{failure}$ = 0.28 vs. $M_{success}$ =0.23; *t (99)* = 3.943; *p* < 0.001); ($M_{failure}$ = -0.11 vs. $M_{success}$ = -0.01; *t (99)* = 3.746; *p* < 0.008).

Action Units. Action Unit 43, i.e., the movement of eyes closed, suggesting a defensive mechanism. The Independent-Samples T-Test revealed that participants in the failure condition reported higher movement action unit 43 ($M_{failure}$ = 0.23 vs. $M_{success}$ =0.14; *t (99)* = -3.436; *p* < 0.001).

Behavioral Metric. A logistic regression was used to test if the participants in the success vs. failure conditions liked or disliked the suggestion provided by AI. The results show that participants in the success condition liked the suggestion more than those in the failure condition ($\chi^2(1)$ = 5.066, p= 0.024).

Mediation Effect of Arousal and Self-Identity Connection. A mediation analysis was conducted using Hayes Process (model 6, Hayes, 2017; n=5000). The mediators were self-identity connection and arousal, the independent variable was failure vs. success, and the dependent variable was a behavioral question of like or dislike. The results revealed that success vs. failure significantly affects self-identity connection (b=-0.8805, t=-2.532. p=0.127. and significantly affects arousal (b=-0.0379, t=-3.223, p=0.0017). In addition, it showed that self-identity connection has a significant effect on arousal (b=0.0062, t=1.999, p=0.048). Self-identity connection has a significant effect on

like/dislike (b=-1.7739, t=-2.794, p=0.0052), and arousal has a significant effect on like/dislike (b=-19.2478, t=-2.225, p=0.0261). Finally, the direct effect was eliminated as success vs. failure does not significantly affect like/dislike (b=0.3816, t=0.271, p=0.7865).

The study reveals that AI classification failures can lead to higher negative emotions among consumers (as proxy of regret), as they may feel that AI may not align with their self-identity. This is because their like or dislikes are mediated through self-identity connection and arousal. When streaming users experience AI classification failures, it reduces arousal, lowers self-identity connections, and decreases enjoyment of suggested content. This research highlights the downstream consequences of AI classification failures on consumers.

4. General Discussion

Three experimental studies show that AI classification failure reduces consumers' self-identity connection, leading to negative regret and subsequent behaviors like subscriptions. Self-identity expression influences smart service failure processing, providing theoretical and managerial implications for the Smart SFR account, AI classification experience, and identity-based motivation literature.

4.1 Theoretical and Managerial Implications

Our primary contribution is to research on smart service failures (Grégoire & Mattila, 2021). Our research bridges the research on AI classification experience (Puntoni et al., 2021) and identity-based motivation (Oyserman, 2009) to investigate an explored stream of smart service failure. We explore how identity expression shapes how smart service failures are processed. Naturally, consumers form strong emotional attachments when they indulge in high-identity expression (Bagozzi et al., 2022). We contribute to past research by showing that the self-identity connection that consumers create intensifies the negative effects of AI classification failure. Furthermore, our findings reveal the downstream consequences of an AI classification failure on consumers' sense of self. The study explores the impact of AI classification failure on consumers' self-identity connection and how it influences their responses to smart service failures (Huang & Philp, 2021). In doing so it contributes to work on AI classification experience (Puntoni et al., 2021) by revealing that AI failures can disrupt consumers' integration with AI agents, reducing their connection with smart services and impacting their ability to engage in self-expression. Failures in AI classification can lead to negative consequences, such as reduced subscription behaviors highlighting the importance of understanding the relationship between AI classification and consumer self-concept. Finally, this study further explores how AI classification failures affect the activation of self-defense

mechanisms in consumers, building on (Baumeister & Hastings (2013) research. Our study investigates how unexpected and possibly ego-threatening situations like errors in AI classification might cause these self-defense mechanisms to go into action. In addition, our research findings have significant practical implications for streaming platforms, particularly in the realm of smart SFR. While the managerial implications of this study are extensive, the specific implications for the streaming context have been relatively understudied. This study aims to fill that gap by shedding light on the consequences of smart service failures within various service domains of streaming platforms. Specifically, we show the downstream consequences of a smart service failure, such as switching or facial expressions and self-defensive mechanisms (e.g., closing eyes to avoid watching AI-recommended content). AI classification failures are inevitable in smart service platforms, and strategies to minimize their effects are crucial. Identity expression plays a significant role in shaping the experience of these failures. Service providers should minimize AI classification failures and support consumers who experience them. Platforms prioritizing user identity expression face unique risks, especially in AI classification failures. Therefore, platforms must balance enabling identity expression with minimizing the negative impact on user experience. By recognizing and understanding these dynamics, streaming platforms can proactively design their systems and processes to accommodate user identities better and mitigate the potential fallout from smart service failures. This requires a thoughtful approach that integrates user-centric considerations and effectively manages the expectations and emotions of users in the face of AI classification failures.

4.2 Limitations and Future Research

This research, although insightful, presents a few limitations. The study's focus on AI implementation for content selection on streaming platforms does not address AI-human collaboration's implications, particularly in classification failures. Understanding the consequences of such failures could help develop more robust and effective systems, shedding light on human-AI interaction dynamics. Future studies should also explore a broader range of digital services to better understand the consequences of AI classification failures on self-identity. This will help researchers compare outcomes across different contexts and identify contextual nuances. This will help develop AI systems that better align with user expectations and requirements, addressing limitations and exploring unexplored areas.

References

Adam, M., Wessel, M., & Benlian, A. (2021). AI-based chatbots in customer service and their effects on user compliance. *Electronic Markets*, *31*(2), 427–445.

Bagozzi, R. P., Brady, M. K., & Huang, M.-H. (2022). AI Service and Emotion. In *Journal of Service Research* (Vol. 25, Issue 4, pp. 499–504). SAGE Publications Sage CA: Los Angeles, CA.

Baumeister, R. F., & Hastings, S. (2013). Distortions of collective memory: How groups flatter and deceive themselves. In *Collective memory of political events* (pp. 277–293). Psychology Press.

Choi, S., Mattila, A. S., & Bolton, L. E. (2020). To Err Is Human(-oid): How Do Consumers React to Robot Service Failure and Recovery? *Journal of Service Research*, *24*(3), 354–371.

Czerwonka, M., & Karwowski, M. (2018). The order matters: Asking about creative activity calibrates creative self-concept. *Creativity Research Journal*, *30*(2), 179–186.

Grégoire, Y., & Mattila, A. S. (2021). Service failure and recovery at the crossroads: recommendations to revitalize the field and its influence. In *Journal of Service Research* (Vol. 24, Issue 3, pp. 323–328). SAGE Publications Sage CA: Los Angeles, CA.

Huh, J., Kim, H. Y., & Lee, G. (2023). "Oh, happy day!" Examining the role of AI-powered voice assistants as a positive technology in the formation of brand loyalty. Journal of Research in Interactive Marketing.

Kirmani, A. (2009). The self and the brand. *Journal of Consumer Psychology*, *19*(3), 271–275.

Murphy, M. C., & Dweck, C. S. (2016). Mindsets shape consumer behavior. *Journal of Consumer Psychology*, *26*(1), 127–136.

Oyserman, D. (2009). Identity-based motivation and consumer behavior. *Journal of Consumer Psychology*, *19*(3), 276–279.

Patterson, P. G., Cowley, E., & Prasongsukarn, K. (2006). Service failure recovery: The moderating impact of individual-level cultural value orientation on perceptions of justice. *International Journal of Research in Marketing*, *23*(3), 263–277.

Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and Artificial Intelligence: An Experiential Perspective. *Journal of Marketing*, *85*(1), 131–151.

Reed II, A., Forehand, M. R., Puntoni, S., & Warlop, L. (2012). Identity-based consumer behavior. *International Journal of Research in Marketing*, *29*(4), 310–321.

Tsiros, M., & Mittal, V. (2000). Regret: A Model of Its Antecedents and Consequences in Consumer Decision Making. *Journal of Consumer Research*, *26*(4), 401–417.