Too Personal to Persuade: Consumer Reactance and the Inferential Personalization Paradox in Responsible Generative AI Marketing

Muhammad Safizal Abdullah[1] , Wan Mashumi Wan Mustafa[2] , Muhamad Azan Yaakob[3] ,Nurfareena Zahari[4] , Muhammad Asyraf Mohd Kassim[5]

Faculty of Business & Communication, Universiti Malaysia Perlis (UniMAP), Uniciti Alam Campus, 02100 Padang Besar, Perlis, Malaysia

DOI: https://doi.org/10.47772/IJRISS.2026.100800471

Received: 26 August 2026; Accepted: 31 August 2026; Published: 08 September 2026

ABSTRACT

Generative artificial intelligence (GenAI) is compressing the distance between algorithmic inference and personalized persuasion. Consumer signals can be interpreted, synthesized into latent knowledge, and converted into bespoke messages within the same interaction. This capability can improve relevance while simultaneously making a firm's inferential reach visible. This paper develops the Inferential Personalization Paradox (IPP), in which the capabilities that enhance personalization value also increase the risk that consumers judge algorithmically constructed knowledge as illegitimately knowable or actionable. Integrating Psychological Reactance Theory with Contextual Integrity, the paper conceptualizes perceived inferential boundary violation (PIBV) as the legitimacy appraisal connecting inferential personalization to autonomy threat and reactance. PIBV comprises inferential knowability and inferential actionability. Inferential salience activates the appraisal process, while contextual congruence and meaningful inferential agency define conditions under which personalization can remain autonomy-preserving. Eight propositions specify a dual-pathway framework and a shift from compensatory privacy calculus toward autonomy defense as PIBV intensifies. The framework advances consumer psychology and socially responsible marketing by defining responsible personalization as legitimate, contestable use of inferred knowledge rather than maximal predictive accuracy.

Keywords: generative AI; personalization; psychological reactance; consumer autonomy; privacy; socially responsible marketing; algorithmic inference

INTRODUCTION

Artificial intelligence has made an old marketing ambition newly consequential: understanding an individual well enough to produce the right intervention at the right moment. Earlier personalization systems used segments, transaction histories, recommender models, and behavioral targeting to select among pre-existing alternatives. Contemporary generative AI (GenAI) adds a different capability. It can translate inferred preferences, goals, and situational states directly into newly generated text, imagery, recommendations, or conversational responses. The theoretically important shift is therefore not that AI suddenly became capable of inference. Predictive systems inferred consumer characteristics long before GenAI (Davenport et al., 2020; Huang & Rust, 2021; Matz et al., 2020). Rather, GenAI compresses the distance between inference and intervention, allowing inferred knowledge to become an immediate ingredient in persuasion (Cillo & Rubera, 2025; Grewal et al., 2025; Matz et al., 2024).

This compression creates a distinctive consumer experience. A retailer remembering a prior purchase can signal ordinary relationship memory. A system appearing to infer financial strain, emotional vulnerability, pregnancy, loneliness, grief, or an impending life transition can signal something else: that the firm has constructed personal knowledge that the consumer did not knowingly provide. The message can be highly relevant and still feel illegitimate. Accuracy itself can become evidence of informational reach.

This problem is increasingly central to socially responsible marketing. Responsible marketing requires more than technically lawful data use or accurate recommendations. It requires design choices that protect consumer

well-being, agency, dignity, and trust while firms pursue relevance and performance (Ferrell & Ferrell, 2022; Kumar et al., 2025). Recent consumer research likewise shows that algorithms can constrain consumer experience, that AI-enhanced personalization can weaken agency, and that autonomy-preserving safeguards matter to how consumers evaluate AI-mediated markets (Hsu, 2026; Kaliyamurthy & Schau, 2025; Karichalil & Kaul, 2026; Tiwari, 2026). The emerging challenge is therefore not simply how to personalize more effectively, but how to distinguish useful personalization from inferential overreach.

Existing marketing research provides important pieces of this puzzle. Personalization can improve usefulness and response while also increasing vulnerability, intrusiveness, irritation, privacy concern, and avoidance (Aguirre et al., 2015; Baek & Morimoto, 2012; Bleier & Eisenbeiss, 2015; White et al., 2008). Broader privacy scholarship shows that consumer responses reflect not only security risk but also behavioral advertising practices, perceived vulnerability, and the legitimacy of data use (Acquisti et al., 2015; Boerman et al., 2017; Martin & Murphy, 2017). Privacy calculus explains why consumers tolerate data-related risk in exchange for valued benefits (Dinev & Hart, 2006), while the privacy paradox documents gaps between stated concerns and behavior (Barth & de Jong, 2017; Kokolakis, 2017). Contextual Integrity explains why information flows are judged relative to contextual norms (Nissenbaum, 2004, 2010). Communication Privacy Management and information-boundary research show that personalized advertising can violate boundaries governing information source, platform, purpose, time, or co-ownership (Petronio, 2002; Sutanto et al., 2013; Zhu et al., 2023).

Recent AI-personalization studies also demonstrate that personalization can create non-linear utility-threat trade-offs, creepiness, surveillance perceptions, reduced autonomy, and reactance (Petrova et al., 2026; Sun et al., 2025; Tangari et al., 2026; Tiwari, 2026; Ul Haq et al., 2026). These findings make the general claim that "too much personalization backfires" insufficient. The unresolved issue is the object of the legitimacy judgment. Prior boundary constructs mainly concern existing information that is collected, transferred, disclosed, or used across an accepted boundary. AI systems can additionally synthesize multiple signals into a new latent conclusion. The perceived violation can therefore reside in the act of algorithmic knowledge construction and in the decision to make that knowledge commercially actionable.

This paper argues when and why does real-time generative personalization shift from useful consumer understanding to illegitimate inferential overreach, thereby activating autonomy defense and psychological reactance? It develops the Inferential Personalization Paradox (IPP) and introduces perceived inferential boundary violation (PIBV) as the central legitimacy appraisal. PIBV concerns whether AI-constructed knowledge is legitimately knowable and legitimately actionable in a marketing relationship. Psychological Reactance Theory (PRT) explains why an inferential violation can become a motivational threat to freedom (Brehm, 1966; Brehm & Brehm, 1981), while Contextual Integrity explains why the same technical inference can be acceptable in one relationship and inappropriate in another (Nissenbaum, 2004, 2010).

The paper makes four contributions. First, it conceptualizes inference-to-intervention compression as the GenAI-specific mechanism that makes inferential reach more immediate and psychologically diagnostic. Second, it distinguishes PIBV from existing privacy-boundary constructs by focusing on the legitimacy of algorithmically constructed knowledge, with separable judgments of inferential knowability and actionability. Third, it identifies inferential salience as an activation condition and contextual congruence and meaningful inferential agency as conditions for autonomy-preserving personalization. Fourth, it theorizes a falsifiable shift from compensatory privacy calculus to autonomy defense, explaining why personalization value can lose its capacity to compensate once inferential entitlement becomes the focal issue. Together, these contributions reposition responsible AI personalization from a question of how accurately firms can know consumers to a question of what firms are legitimately entitled to infer, use, and make contestable.

THEORETICAL FOUNDATION: FROM PERSONALIZATION VALUE TO INFERENTIAL LEGITIMACY

Inference-to-intervention compression

This paper defines real-time generative AI hyper-personalization as the dynamic generation or adaptation of marketing content using current and historical consumer signals, including inferred states, such that the persuasive output can change during or near the moment of interaction. Its distinctive feature is inference-to-intervention compression: latent conclusions that once informed an intermediate score, segment, or recommendation can now shape the persuasive object itself.

Four inference-revealing affordances determine how strongly this capability can expose algorithmic knowing. Inferential intimacy concerns the sensitivity or psychological depth of the inferred state. Inferring shoe preference differs from inferring grief, financial distress, or emotional vulnerability. Temporal immediacy concerns how quickly a signal becomes an intervention; a message appearing seconds after a revealing behavior can make monitoring and inference salient. Cross-context synthesis concerns combining traces from contexts consumers mentally separate. Generative specificity concerns how uniquely the resulting message appears fitted to a consumer's inferred condition. These dimensions are not interchangeable and should be treated as a formative configuration of inferential intensity rather than as reflective manifestations of a single latent trait.

The framework does not assume that inference is inherently harmful or that AI is objectively omniscient. Recommendation systems routinely infer preferences that consumers welcome, and a recent meta-analysis suggests that the effectiveness of psychological targeting is heterogeneous rather than uniformly powerful (Perla et al., 2026). The mechanism proposed here is therefore perceptual and normative: what matters is whether the consumer experiences the inferred knowledge as appropriate for the relationship and purpose. This is especially important because consumers can withhold an explicit disclosure and remain inferable from other traces (Matz et al., 2020; Schoenebeck et al., 2024). Privacy therefore extends beyond controlling input data toward judging what conclusions firms may legitimately construct from those data. This experiential emphasis is consistent with broader work on how AI changes consumer experiences of classification, delegation, and data capture (Puntoni et al., 2021).

The Inferential Personalization Paradox

The Inferential Personalization Paradox (IPP) is the condition in which the same inferential capability that increases personalization value also increases the likelihood that consumers experience the resulting knowledge as an illegitimate informational and autonomy boundary crossing. The paradox is strongest when the message is simultaneously useful and objectionably revealing.

This formulation differs from the classic privacy paradox, which concerns inconsistencies between privacy attitudes and behavior, and from broader personalization-privacy trade-offs, which often model privacy as a cost that can be compensated by relevance. The IPP centers on inferential legitimacy. Some practices may be evaluated not as an additional privacy price but as a rule violation: "you should not know this about me" or "even if you know it, you should not use it to persuade me." This distinction creates the possibility of a decision-mode transition from calculating net value to restoring threatened autonomy.

Integrating Contextual Integrity and Psychological Reactance Theory

Contextual Integrity and PRT explain different stages of this process. Contextual Integrity proposes that privacy is preserved when information flows conform to norms concerning actors, information types, purposes, and transmission principles (Nissenbaum, 2004, 2010), an approach increasingly relevant to data-based marketing and innovation (Bleier et al., 2020). For GenAI marketing, the challenge extends from information flow to knowledge synthesis. Several individually acceptable signals can be combined into a conclusion that would itself have been inappropriate to collect directly. The relevant legitimacy questions become: should this marketer be entitled to derive this inference, and should it be entitled to use the inference for this persuasive purpose?

PRT explains what happens when inferential illegitimacy is experienced as a freedom threat. Reactance is a motivational state triggered when a valued freedom is threatened or eliminated (Brehm, 1966; Brehm & Brehm, 1981). In persuasion research, state reactance is commonly represented by intertwined negative cognitions and anger (Dillard & Shen, 2005; Rains, 2013). The present framework extends that logic to an epistemic form of overreach: consumers can resist not because an advertisement explicitly restricts choice, but because the firm appears to appropriate control over which aspects of the self are inferred and activated for persuasion.

Privacy calculus is retained as the baseline decision logic rather than a third theoretical anchor. At low levels of inferential concern, consumers may weigh value against privacy costs. PRT becomes increasingly important when the consumer interprets the inferential act as a threat to autonomy that should be resisted rather than priced.

CONCEPTUAL ARCHITECTURE AND BOUNDARIES

Perceived inferential boundary violation

Perceived inferential boundary violation (PIBV) is the consumer's episode-specific judgment that an AI-enabled marketer has exceeded its legitimate inferential entitlement by constructing or commercially operationalizing personal knowledge beyond what is appropriately knowable or actionable in the relationship.

PIBV contains two related judgments. Inferential knowability concerns whether the marketer should legitimately be able to derive the knowledge. Inferential actionability concerns whether the marketer should be entitled to use that knowledge for the particular persuasive purpose. The distinction matters because derivation and use can diverge. A financial service may be expected to infer cash-flow stress for fraud prevention while consumers reject use of the same inference to intensify a high-interest credit appeal.

PIBV differs from prior information-boundary intrusion. Existing boundary research examines whether information crosses accepted boundaries of disclosure, platform, purpose, time, or ownership (Sutanto et al., 2013; Zhu and Kanjanamekanant, 2021; Zhu et al., 2023). PIBV evaluates a transformation from available signals into a newly constructed representation of the person, followed by the commercial use of that representation. It is also distinct from general privacy concern, surveillance, intrusiveness, creepiness, manipulation, autonomy threat, and reactance.

Table 1. Conceptual position of PIBV relative to adjacent constructs

Construct	Core question	Distinction from PIBV
Privacy concern	Are my data vulnerable, misused, or outside my control?	Broad orientation; can exist without a specific inference.
Perceived surveillance	Am I being watched or tracked?	Focuses on observation, not the legitimacy of what the system concludes.
Intrusiveness	Does this communication interfere with my experience or personal space?	A message can be intrusive without relying on an illegitimate inference.
Creepiness	Does the encounter feel unsettling, ambiguous, or eerie?	Broader affective episode; PIBV is a normative-cognitive legitimacy appraisal.
Perceived manipulation	Is the firm covertly steering or exploiting me?	Focuses on persuasive intent rather than entitlement to infer and use knowledge.

Autonomy threat	Is my freedom or control being constrained?	Downstream meaning of PIBV, not the inferential legitimacy judgment itself.
Psychological reactance	Am I motivationally resisting a threatened freedom?	Motivational response after autonomy threat.

Inferential salience, contextual congruence, and meaningful inferential agency

A consumer cannot judge inferential entitlement without some basis for recognizing that inference occurred. Inferential salience is therefore an activation condition: the degree to which the output provides cues that the marketer has derived and used latent personal knowledge. Salience can arise from sensitive content, unusual timing, cross-context fit, generative specificity, or explicit explanations. This condition explains why extensive personalization often remains psychologically uneventful. If inferential reach is not noticed or suspected, PIBV may never become focal. Tracking-awareness and ad-transparency research supports the importance of such recognition (Kim et al., 2019; Tangari et al., 2026).

Contextual congruence is the perceived fit between the origin and meaning of the information, the inference drawn from it, and the purpose for which the inference is used. High congruence legitimizes some intensive personalization because the consumer views the inference as an expected part of the relationship. Low congruence makes the same inferential capability more likely to be judged as overreach.

Meaningful inferential agency is the consumer's perceived ability to contest, correct, restrict, revoke, or redirect the use of inferred knowledge in ways that materially change subsequent personalization. It differs from transparency. Transparency supplies information; agency supplies consequential recourse. An explanation can increase inferential salience without restoring freedom. By contrast, the ability to inspect an inferred category, correct it, prevent its use, or revoke it can reduce the autonomy threat generated by a boundary violation (Dietvorst et al., 2018; Fink et al., 2024; Tucker, 2014).

Competing explanations and the explanatory role of PIBV

The framework overlaps with several established explanations of resistance to personalized persuasion, but it addresses a different theoretical question. Creepiness captures an aversive emotional episode associated with ambiguity, intrusive surveillance, or unexpected technological behavior (Petrova et al., 2026). A consumer may find an AI interface creepy without believing that it has constructed illegitimate knowledge, and an unobtrusive interface may still trigger PIBV if a highly accurate message reveals an unacceptable inference. Creepiness can therefore accompany PIBV without substituting for it.

Persuasion Knowledge explains how consumers recognize, interpret, and cope with a marketer's influence tactics (Friestad & Wright, 1994). It becomes especially relevant when hyper-personalization signals strategic exploitation or manipulative intent. PIBV is conceptually prior to that judgment in a narrower sense: it asks whether the knowledge used as persuasive input should have been derived or activated at all. A consumer can recognize a persuasion tactic while accepting the underlying inference, or reject the inferential entitlement even when the persuasive tactic itself is transparent and non-deceptive.

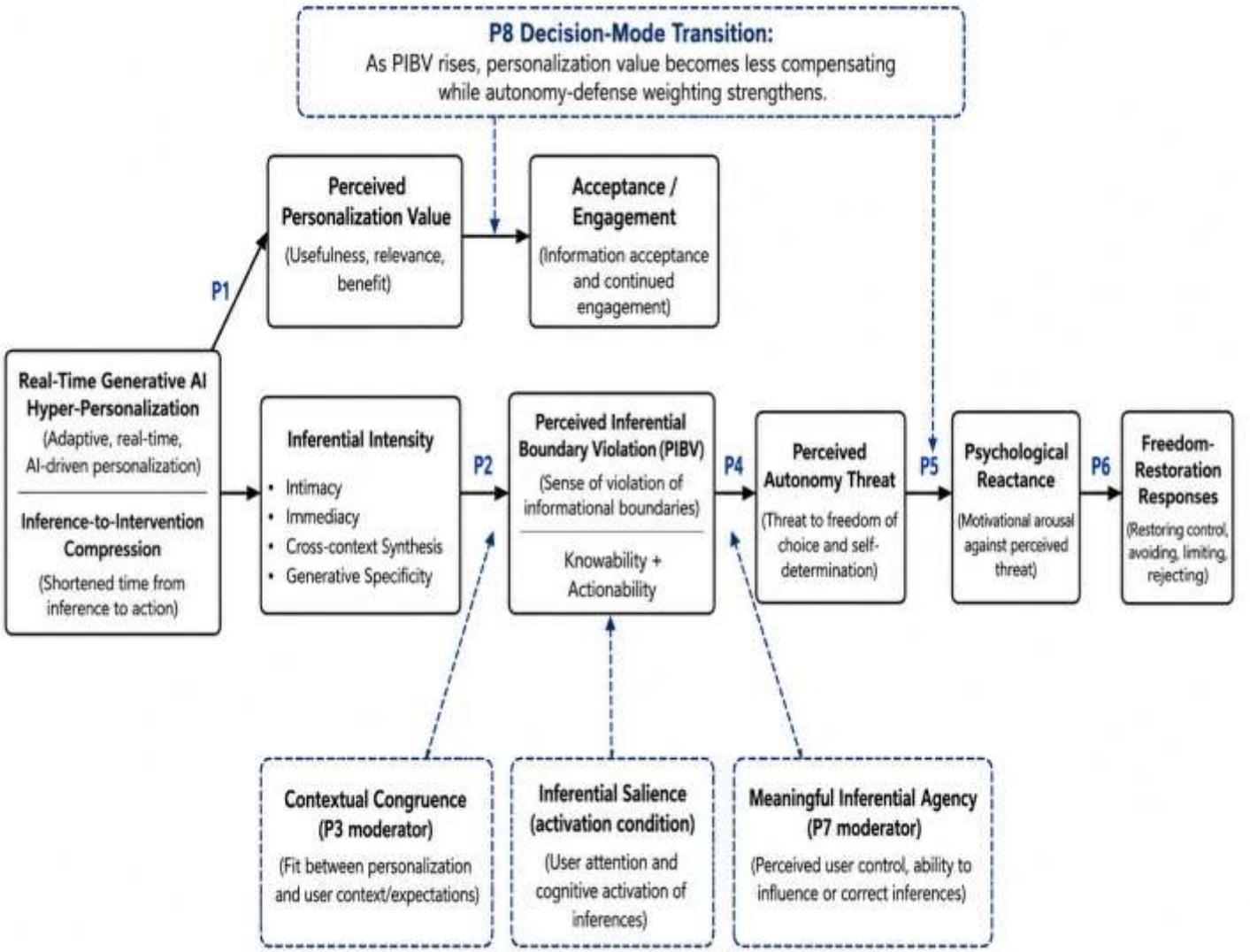
Algorithm aversion concerns reluctance to rely on algorithmic judgment, often because of error, task mismatch, or reduced control (Dietvorst et al., 2018). PIBV does not require distrust in algorithmic competence. Accuracy may intensify the problem because a correct inference can reveal how much the system appears to know. Likewise, privacy calculus explains continuing trade-offs between personalization benefits and privacy costs, whereas PIBV identifies the legitimacy appraisal that can make such trade-offs increasingly non-compensatory.

These distinctions define the explanatory role of PIBV. It does not claim that inferential legitimacy is the only route to AI resistance. Rather, it specifies a route that adjacent perspectives do not isolate: resistance arising because algorithmically constructed knowledge is judged to exceed the marketer's legitimate entitlement to know or act. This narrower mechanism permits competing explanations to be modeled as alternatives, antecedents, or

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VIII August 2026

parallel responses rather than being absorbed into a generic negative-reaction construct.4. Conceptual Framework and Propositions

Figure 1. The Inferential Personalization Paradox: value creation, autonomy threat, and the calculus-to-autonomy-defense transition



The framework contains two simultaneous pathways. The value pathway explains why firms and consumers benefit from personalization. The autonomy-threat pathway explains when the same inferential capability becomes psychologically unacceptable. Inferential salience is required for a legitimacy appraisal, contextual congruence conditions the transition from inferential intensity to PIBV, and meaningful inferential agency conditions the transition from PIBV to autonomy threat. P8 integrates the pathways by theorizing a change in the consumer's decision mode.

Personalization value

Personalization creates value when it reduces mismatch between the consumer's current goal and the firm's response. GenAI can increase that value because it can tailor not only which offer is shown but how the offer is framed, explained, sequenced, and conversationally adapted. Experimental evidence shows that generative models can produce psychologically personalized persuasion at scale, while AI-personalization research continues to document relevance and usefulness benefits (Matz et al., 2024; Tangari et al., 2026). The beneficial pathway is necessary to the IPP because rejection is not paradoxical when the personalized intervention has no value.

P1. Real-time generative AI hyper-personalization increases perceived personalization value when generated content is interpreted as relevant to consumers' active goals and situational needs.

Inferential intensity and PIBV

The threat pathway begins when personalization reveals the system's inferential reach. Intimate inferences are more personally consequential; immediate interventions link the consumer's behavior to persuasion tightly enough to make inference visible; cross-context synthesis violates mental separations between domains; and generative specificity makes the message diagnostic of what the system appears to know. Sensitive-inference research shows that comfort depends on more than relevance, while consumers evaluate inferred information less favorably than information they explicitly supplied (Kim et al., 2019; Schoenebeck et al., 2024). These affordances should therefore increase the probability of a PIBV appraisal when they become sufficiently salient.

P2. Greater inferential intensity, formed by inferential intimacy, temporal immediacy, cross-context synthesis, and generative specificity, increases perceived inferential boundary violation when the inferential basis of personalization is sufficiently salient.

Contextual congruence

The same inference can be legitimate in one relationship and unacceptable in another. A fitness service can infer exercise preferences from workout activity without necessarily violating expectations, whereas transferring the same traces into unrelated financial or emotional targeting creates a different normative meaning. Contextual Integrity predicts that consumers evaluate information practices relative to the purposes and roles of the context, and evidence shows less favorable responses when advertising relies on information obtained outside the immediate context (Hanson et al., 2020; Kim et al., 2019).

P3. Contextual congruence weakens the positive relationship between inferential intensity and perceived inferential boundary violation, such that equivalent inferences are judged as less boundary-violating when their derivation and use fit the norms of the consumer-marketer relationship.

PIBV and autonomy threat

A judgment of inferential illegitimacy becomes motivationally consequential when it changes the consumer's perceived control over self-definition and persuasion. The consumer can object even when the inference is accurate and no security breach occurred. What is threatened is the freedom to decide whether a personal state should become commercially actionable. This is especially salient when the system can categorize and act on the person faster than the person can understand, contest, or revise the categorization.

P4. Perceived inferential boundary violation increases perceived autonomy threat by signaling that the marketer has constructed or operationalized personal knowledge beyond the consumer's perceived inferential entitlement.

Autonomy threat and psychological reactance

PRT predicts that a salient threat to a valued freedom produces reactance (Brehm, 1966). In the present context, the threatened freedom concerns control over whether personal identities, vulnerabilities, or circumstances are used as persuasive inputs. Reactance should therefore involve anger and negative cognitions such as counterarguing, motive suspicion, rejection of the persuasive attempt, and unfavorable judgments of the brand's entitlement to influence (Dillard & Shen, 2005; Rains, 2013). Recent AI research similarly places autonomy at the center of consumer resistance to personalization (Hsu, 2026; Karichalil & Kaul, 2026; Tiwari, 2026).

P5. Perceived autonomy threat generated by inferential boundary violation increases psychological reactance, manifested through heightened anger and negative cognitions toward the AI-mediated persuasion attempt.

Freedom restoration

Reactance motivates restoration rather than determining a single behavior. Consumers can ignore or counterargue a message, restrict permissions, reject personalized features, delete histories, withhold or falsify information, correct an inference, switch brands, complain, or exit the relationship. The specific response should depend on the freedom that appears threatened. Misclassification can invite correction; unwanted tracking can invite technical restriction; use of a sensitive but accurate inference can motivate revocation or relationship exit. This distinction is important because reactance is the motivational state, while avoidance, opacity, and switching are downstream strategies. Repeated restoration episodes can eventually reduce brand trust and relationship continuity (Baek & Morimoto, 2012; Petrova et al., 2026).

P6. Psychological reactance increases freedom-restoration responses that reduce the marketer's persuasive or inferential reach; repeated or relationship-level restoration responses subsequently increase the likelihood of trust erosion and relationship withdrawal.

Meaningful inferential agency

Responsible personalization requires more than a disclosure that an algorithm is operating. Consumers need consequential routes to contest the representation being used. Consumer agency matters when marketing systems define or categorize the self, while privacy controls and even limited opportunities to modify algorithmic output can improve acceptance (Bhattacharjee et al., 2014; Dietvorst et al., 2018; Fink et al., 2024; Martin et al., 2017; Tucker, 2014). The present framework predicts that agency is most important after a legitimacy violation has been perceived. It need not erase the judgment that the firm overreached, but it can prevent that judgment from becoming a severe freedom threat by making restoration possible within the relationship.

P7. Meaningful inferential agency weakens the positive relationship between perceived inferential boundary violation and autonomy threat, thereby reducing the indirect effect of PIBV on psychological reactance.

From privacy calculus to autonomy defense

The most distinctive proposition concerns the decision rule connecting the value and threat pathways. Privacy calculus assumes that consumers compare expected personalization benefits with privacy-related costs (Dinev & Hart, 2006), and contemporary marketing research continues to formalize privacy-utility trade-offs as an optimization problem (Ponte et al., 2024). This logic explains why people continue to use data-intensive services despite concern. It also implies that sufficiently high value can compensate for sufficiently high privacy cost.

PIBV creates a different possibility. Some inferential practices can become **non-compensatory** because they are treated as violations of a valued rule or protected freedom rather than as larger costs inside the same utility function. Research on protected values and taboo trade-offs shows that people sometimes resist making compensatory exchanges when a decision implicates a norm they regard as non-tradeable (Baron & Spranca, 1997; Fiske & Tetlock, 1997). Applied to AI personalization, the consumer may stop asking whether a relevant offer is "worth" the privacy cost and instead ask how to restore control over an illegitimate use of the self.

This claim is falsifiable. A simple high-cost model predicts that personalization value should continue to increase acceptance even when privacy costs are very high, provided the benefit becomes sufficiently large. A decision-mode transition predicts something different: as PIBV increases, the marginal effect of personalization value on acceptance should weaken, while the effect of autonomy threat on freedom-restoration behavior should strengthen. Rejection of a useful offer can itself become instrumentally valuable because it reasserts autonomy. The strongest form of the paradox occurs when greater relevance increases perceived overreach because the quality of personalization demonstrates how much the system has inferred.

P8. As perceived inferential boundary violation intensifies, consumer evaluation shifts from compensatory personalization-privacy calculus toward autonomy defense, such that the positive effect

of personalization value on acceptance progressively weakens while the influence of autonomy threat on freedom-restoration responses strengthens.

DISCUSSION AND THEORETICAL CONTRIBUTIONS

The framework advances consumer psychology by relocating the problem of AI personalization from data quantity to inferential entitlement. GenAI is distinctive not because it invented consumer inference but because it compresses the sequence from signal to latent conclusion to persuasive intervention. This compression makes an invisible modeling process more likely to become experientially visible. The consumer encounters not merely an algorithmic score but a message whose timing and specificity can reveal the score's psychological meaning.

Second, PIBV extends privacy-boundary scholarship from the movement of existing information to the legitimacy of constructed knowledge. This distinction avoids relabeling established concepts. Communication Privacy Management and prior boundary work explain when information crosses boundaries of disclosure, purpose, platform, or co-ownership (Petronio, 2002; Sutanto et al., 2013; Zhu et al., 2023). PIBV addresses a transformation problem: whether a marketer should be entitled to create a particular inference from otherwise available signals and whether it should be entitled to commercialize that inference. The knowability-actionability distinction is especially useful for responsible marketing because it allows governance to restrict persuasive use even when derivation serves a legitimate service or safety function.

Third, the framework extends PRT by conceptualizing epistemic overreach as a precursor to freedom threat. AI persuasion can appear facilitative, polite, and choice-expanding while still triggering resistance if consumers believe that a firm has appropriated a personal state for influence. The threatened freedom therefore concerns not only choice among offers but control over commercial self-definition. Distinguishing reactance from freedom-restoration behavior also produces a testable process rather than treating any negative outcome as reactance.

Fourth, inferential salience explains why powerful personalization often produces no backlash. Consumers cannot object to inferential entitlement that never becomes psychologically visible. This activation condition also clarifies why transparency is not uniformly beneficial. An explanation can reveal an objectionable inference and increase PIBV if the consumer has no recourse. Transparency without agency can expose the violation; agency can restore a route to freedom. This transparency-agency asymmetry shifts responsible design away from notice alone toward contestability.

Finally, P8 extends personalization-privacy research by distinguishing high privacy cost from a change in decision logic. The IPP predicts that value and privacy are compensatory only up to the point at which inferential legitimacy becomes a protected autonomy concern. This provides a psychological mechanism for understanding why highly relevant personalization can fail precisely because it is highly revealing. It also aligns the framework with socially responsible marketing: responsible AI is not merely AI that predicts accurately or discloses its operation, but AI whose inferences are contextually legitimate, proportionate to the relationship, and subject to meaningful consumer agency.

Implications for Socially Responsible AI Marketing

The framework suggests a design principle for responsible personalization: optimize legitimate relevance, not maximum inference. Firms should distinguish what models can infer from what the customer relationship authorizes them to infer and use. An inferential sensitivity map can classify ordinary preference inferences separately from health, financial, relational, identity, and emotional-vulnerability inferences. High-sensitivity inferences should face stronger purpose restrictions and, in some contexts, categorical exclusion from persuasion.

Second, firms should audit inference-to-intervention compression. Real-time activation can create value, but immediate persuasive use of an intimate inference can amplify vulnerability and reactance. Responsible design

may therefore require friction: delayed activation, additional human review, context-specific prohibitions, or a separation between service-support inferences and persuasion-oriented inferences.

Third, meaningful inferential agency should be designed into the consumer experience. Instead of only offering a generic "Why am I seeing this?" explanation, systems can expose high-impact inferred categories and allow consumers to correct, restrict, or revoke them. For example, "We are using an assumption that you are shopping for a new home. Is that correct?" makes the inference contestable. Such mechanisms can improve both dignity and data quality by replacing silent resentment or deliberate falsification with correction.

Fourth, transparency should be conditional rather than treated as a universal remedy. If disclosure reveals an unacceptable inference but offers no means of recourse, it can increase inferential salience and worsen reactance (Kim et al., 2019). Responsible transparency therefore requires actionable intelligibility: explanations linked to controls that change what the system does next.

The broader governance question is simple: not "How much can our AI know?" but "What are we legitimately entitled to infer and use in this relationship, and can the consumer meaningfully contest it?" This reframing connects AI personalization directly to consumer well-being, trust, and socially responsible experience design.

Research Agenda

The first priority is measurement. PIBV should be developed as a higher-order construct with distinguishable knowability and actionability dimensions and tested against privacy concern, surveillance, intrusiveness, creepiness, perceived manipulation, autonomy threat, and reactance. Scale development should include contexts in which the inference is accurate but unwanted so that legitimacy is not confounded with perceived error.

Experimental work should manipulate the four components of inferential intensity independently. Studies can hold recommendation relevance constant while varying the intimacy of the inferred state, the delay between signal and intervention, the contextual origin of signals, and the specificity of generated content. Separate manipulations of objective inference and inferential salience would clarify whether consumers react to what the system actually does or to what the output allows them to infer about the system.

Knowability and actionability should also be manipulated separately. A firm may be permitted to derive a risk signal for service protection but not to use it for commercial persuasion. Such designs can identify whether responsible use restrictions mitigate reactance even when the underlying inference remains available.

The transparency-agency asymmetry is well suited to factorial experiments crossing explanation with meaningful control. The framework predicts that transparency without agency may increase PIBV, whereas transparency combined with correction and revocation should reduce autonomy threat. Behavioral outcomes should include opt-out, inference correction, data withholding, falsification, switching, complaints, and actual choice rather than intentions alone.

Most importantly, researchers should test P8 against a rival continuous-cost model. Experiments can jointly manipulate personalization value and PIBV and estimate whether the marginal effect of value on acceptance collapses as violation intensifies. Piecewise, mixture, or choice models may determine whether responses are better explained by a smooth cost-benefit function or by heterogeneous transitions into autonomy-defense modes. Longitudinal work can then examine whether repeated minor violations accumulate into stable brand distrust or strategic consumer opacity.

CONCLUSION

GenAI makes personalization more powerful by compressing the distance between algorithmic inference and persuasive intervention. The same capability can make personalization psychologically fragile. Consumers may value a message while objecting to the knowledge that made the message possible. The Inferential Personalization Paradox explains this tension through PIBV, autonomy threat, reactance, and freedom restoration, while contextual congruence and meaningful inferential agency identify conditions for responsible

design. The theoretical frontier of AI marketing is therefore not simply predictive accuracy. It is inferential legitimacy: whether knowledge is appropriately constructed, appropriately used, and meaningfully contestable by the consumer.

Data availability statement

Data sharing is not applicable to this conceptual article because no new empirical data were created or analyzed.

Conflict of interest statement

The authors declare no conflict of interest related to this conceptual article.

REFERENCES

1. Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509-514. https://doi.org/10.1126/science.aaa1465
2. Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. Journal of Retailing, 91(1), 34-49. https://doi.org/10.1016/j.jretai.2014.09.005
3. Baek, T. H., & Morimoto, M. (2012). Stay away from me: Examining the determinants of consumer avoidance of personalized advertising. Journal of Advertising, 41(1), 59-76. https://doi.org/10.2753/JOA0091-3367410105
4. Baron, J., & Spranca, M. (1997). Protected values. Organizational Behavior and Human Decision Processes, 70(1), 1-16. https://doi.org/10.1006/obhd.1997.2690
5. Barth, S., & de Jong, M. D. T. (2017). The privacy paradox: Investigating discrepancies between expressed privacy concerns and actual online behavior: A systematic literature review. Telematics and Informatics, 34(7), 1038-1058. https://doi.org/10.1016/j.tele.2017.04.013
6. Bhattacharjee, A., Berger, J., & Menon, G. (2014). When identity marketing backfires: Consumer agency in identity expression. Journal of Consumer Research, 41(2), 294-309. https://doi.org/10.1086/676125
7. Bleier, A., & Eisenbeiss, M. (2015). Personalized online advertising effectiveness: The interplay of what, when, and where. Marketing Science, 34(5), 669-688. https://doi.org/10.1287/mksc.2015.0930
8. Bleier, A., Goldfarb, A., & Tucker, C. (2020). Consumer privacy and the future of data-based innovation and marketing. International Journal of Research in Marketing, 37(3), 466-480. https://doi.org/10.1016/j.ijresmar.2020.03.006
9. Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioral advertising: A literature review and research agenda. Journal of Advertising, 46(3), 363-376. https://doi.org/10.1080/00913367.2017.1339368
10. Brehm, J. W. (1966). A theory of psychological reactance. Academic Press.
11. Brehm, S. S., & Brehm, J. W. (1981). Psychological reactance: A theory of freedom and control. Academic Press.
12. Cillo, P., & Rubera, G. (2025). Generative AI in innovation and marketing processes: A roadmap of research opportunities. Journal of the Academy of Marketing Science, 53, 684-701. https://doi.org/10.1007/s11747-024-01044-7
13. Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. Journal of the Academy of Marketing Science, 48, 24-42. https://doi.org/10.1007/s11747-019-00696-0
14. Dietvorst, B. J., Simmons, J. P., & Massey, C. (2018). Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Management Science, 64(3), 1155-1170. https://doi.org/10.1287/mnsc.2016.2643
15. Dillard, J. P., & Shen, L. (2005). On the nature of reactance and its role in persuasive health communication. Communication Monographs, 72(2), 144-168. https://doi.org/10.1080/03637750500111815
16. Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. Information Systems Research, 17(1), 61-80. https://doi.org/10.1287/isre.1060.0080

17. Ferrell, L., & Ferrell, O. C. (2022). Broadening the definition of socially responsible marketing. Journal of Macromarketing, 42(4), 560-566. https://doi.org/10.1177/02761467221094570
18. Fink, L., Newman, L., & Haran, U. (2024). Let me decide: Increasing user autonomy increases recommendation acceptance. Computers in Human Behavior, 156, 108244. https://doi.org/10.1016/j.chb.2024.108244
19. Fiske, A. P., & Tetlock, P. E. (1997). Taboo trade-offs: Reactions to transactions that transgress the spheres of justice. Political Psychology, 18(2), 255-297. https://doi.org/10.1111/0162-895X.00058
20. Friestad, M., & Wright, P. (1994). The Persuasion Knowledge Model: How people cope with persuasion attempts. Journal of Consumer Research, 21(1), 1-31. https://doi.org/10.1086/209380
21. Grewal, D., Satornino, C. B., Davenport, T., & Guha, A. (2025). How generative AI is shaping the future of marketing. Journal of the Academy of Marketing Science, 53, 702-722. https://doi.org/10.1007/s11747-024-01064-3
22. Hanson, J., Wei, M., Veys, S., Kugler, M., Strahilevitz, L., & Ur, B. (2020). Taking data out of context to hyper-personalize ads: Crowdworkers' privacy perceptions and decisions to disclose private information. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-13). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376415
23. Hsu, Y.-M. (2026). From capability to care: Sense-breaking, sense-giving, and strategic flexibility as drivers of ethical, autonomy-preserving AI personalization. Psychology & Marketing, 43(6), 1418-1431. https://doi.org/10.1002/mar.70118
24. Huang, M.-H., & Rust, R. T. (2021). A strategic framework for artificial intelligence in marketing. Journal of the Academy of Marketing Science, 49, 30-50. https://doi.org/10.1007/s11747-020-00749-9
25. Kaliyamurthy, A. K., & Schau, H. J. (2025). How algorithms constrain consumer experience. Journal of Consumer Research, 52(4), 826-847. https://doi.org/10.1093/jcr/ucaf016
26. Karichalil, R. A., & Kaul, D. (2026). The emancipation paradox: When AI-enhanced personalization undermines consumer agency and what socially responsible marketing requires. Journal of Macromarketing. Advance online publication. https://doi.org/10.1177/02761467261447188
27. Kim, T., Barasz, K., & John, L. K. (2019). Why am I seeing this ad? The effect of ad transparency on ad effectiveness. Journal of Consumer Research, 45(5), 906-932. https://doi.org/10.1093/jcr/ucy039
28. Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Computers & Security, 64, 122-134. https://doi.org/10.1016/j.cose.2015.07.002
29. Kumar, V., Hollebeek, L. D., Sharma, A., Rajan, B., & Srivastava, R. K. (2025). Responsible stakeholder engagement marketing. Journal of Business Research, 189, 115143. https://doi.org/10.1016/j.jbusres.2024.115143
30. Martin, K. D., Borah, A., & Palmatier, R. W. (2017). Data privacy: Effects on customer and firm performance. Journal of Marketing, 81(1), 36-58. https://doi.org/10.1509/jm.15.0497
31. Martin, K. D., & Murphy, P. E. (2017). The role of data privacy in marketing. Journal of the Academy of Marketing Science, 45, 135-155. https://doi.org/10.1007/s11747-016-0495-4
32. Matz, S. C., Appel, R. E., & Kosinski, M. (2020). Privacy in the age of psychological targeting. Current Opinion in Psychology, 31, 116-121. https://doi.org/10.1016/j.copsyc.2019.08.010
33. Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. Scientific Reports, 14, 4692. https://doi.org/10.1038/s41598-024-53755-0
34. Nissenbaum, H. (2004). Privacy as contextual integrity. Washington Law Review, 79, 119-158.
35. Nissenbaum, H. (2010). Privacy in context: Technology, policy, and the integrity of social life. Stanford University Press.
36. Perla, R., Maran, T., Bagci, B., Kraus, S., Kanbach, D. K., & Bouncken, R. B. (2026). The (in)effectiveness of psychological targeting: A meta-analytic review. Psychology & Marketing, 43(3), 575-590. https://doi.org/10.1002/mar.70073
37. Petronio, S. (2002). Boundaries of privacy: Dialectics of disclosure. State University of New York Press.
38. Petrova, A., Malär, L., Hoyer, W., & Krohmer, H. (2026). The phenomenon of creepiness in a digital marketing world. Psychology & Marketing, 43, 834-851. https://doi.org/10.1002/mar.70089

39. Ponte, G. R., Wieringa, J. E., Boot, T., & Verhoef, P. C. (2024). Where's Waldo? A framework for quantifying the privacy-utility trade-off in marketing applications. International Journal of Research in Marketing, 41(3), 529-546. https://doi.org/10.1016/j.ijresmar.2024.05.003
40. Puntoni, S., Reczek, R. W., Giesler, M., & Botti, S. (2021). Consumers and artificial intelligence: An experiential perspective. Journal of Marketing, 85(1), 131-151. https://doi.org/10.1177/0022242920953847
41. Rains, S. A. (2013). The nature of psychological reactance revisited: A meta-analytic review. Human Communication Research, 39(1), 47-73. https://doi.org/10.1111/j.1468-2958.2012.01443.x
42. Schoenebeck, S., Goray, C., Vadapalli, A., & Andalibi, N. (2024). Sensitive inferences in targeted advertising. Northwestern Journal of Technology and Intellectual Property, 21(2), Article 1.
43. Sun, H., Xie, P., & Sun, Y. (2025). The inverted U-shaped effect of personalization on consumer attitudes in AI-generated ads: Striking the right balance between utility and threat. Journal of Advertising Research, 65(2), 237-258. https://doi.org/10.1080/00218499.2025.2462405
44. Sutanto, J., Palme, E., Tan, C.-H., & Phang, C. W. (2013). Addressing the personalization-privacy paradox: An empirical assessment from a field experiment on smartphone users. MIS Quarterly, 37(4), 1141-1164.
45. Tangari, A. H., Bui, M., Krishen, A. S., & Kemp, E. (2026). AI-driven personalization in social media advertising: How tracking awareness and eWOM homophily shape trust and consumer responses. Psychology & Marketing, 43(9), 2175-2192. https://doi.org/10.1002/mar.70166
46. Tiwari, V. (2026). Caught between convenience and autonomy: Psychological dilemmas in consumers' use of AI-enabled personalization. Psychology & Marketing. Advance online publication. https://doi.org/10.1002/mar.70209
47. Tucker, C. E. (2014). Social networks, personalized advertising, and privacy controls. Journal of Marketing Research, 51(5), 546-562. https://doi.org/10.1509/jmr.10.0355
48. Ul Haq, I., Mazhar, B., Tareen, H. K., Maqsood, F., & Wel Henage, U. A. P. (2026). The personalization paradox in AI-driven social media advertising: A tripartite framework. International Journal of Advertising. Advance online publication. https://doi.org/10.1080/02650487.2026.2704450
49. White, T. B., Zahay, D. L., Thorbjørnsen, H., & Shavitt, S. (2008). Getting too personal: Reactance to highly personalized email solicitations. Marketing Letters, 19(1), 39-50. https://doi.org/10.1007/s11002-007-9027-9
50. Zhu, Y.-Q., & Kanjanamekanant, K. (2021). No trespassing: Exploring privacy boundaries in personalized advertisement and its effects on ad attitude and purchase intentions on social media. Information & Management, 58(2), 103314. https://doi.org/10.1016/j.im.2020.103314
51. Zhu, Y.-Q., Kanjanamekanant, K., & Chiu, Y.-T. (2023). Reconciling the personalization-privacy paradox: Exploring privacy boundaries in online personalized advertising. Journal of the Association for Information Systems, 24(1), 294-316. https://doi.org/10.17705/1jais.00775