

Can LLM-Based Digital Twins Approximate Human Loss Aversion and Reward Anticipation in Consumer Decision-Making? A Narrative Review of Neural Benchmarks and Computational Limitations

Sanat Tibrewal¹, Pranesh Nagarajan¹,

¹ Allen High School, Allen, Texas, United States

Corresponding author: Sanat Tibrewal, tibrewalsanat@gmail.com

August 2026

Abstract

Background: Consumer neuroscience has identified loss aversion and reward anticipation as two core mechanisms shaping purchase decisions, but traditional measurement methods such as fMRI and EEG remain expensive and difficult to scale.

Methods: This narrative review synthesizes peer-reviewed literature on prospect theory, dopaminergic reward circuitry, consumer neuroscience, and recent work on large language model (LLM) based digital twins and silicon sampling to evaluate whether LLMs can approximate these mechanisms.

Results: One recent consumer digital-twin study reported 86% future-purchase prediction accuracy using individualized fine-tuning on Amazon transaction data, and a large-scale behavioral benchmark found that even the best-performing LLMs achieved only modest simulation fidelity (40.80/100) across 20 diverse human-behavior datasets. Available studies

have not established that LLMs replicate the behavioral asymmetry of loss aversion, the temporal dynamics of reward anticipation, or individual variation in these mechanisms.

Conclusion: Current evidence suggests that LLMs function as useful behavioral prediction tools for consumer research but have not yet been shown to reproduce the mechanistic processes documented in consumer neuroscience.

Keywords: consumer neuroscience; loss aversion; reward anticipation; digital twins; silicon sampling; neuromarketing; prospect theory; generative AI

1. Introduction

Traditional models of consumer decision-making assume rational agents who maximize utility through deliberate calculation, but empirical evidence shows that systematic deviations from rationality are driven by psychological and neural mechanisms. Neuromarketing and consumer neuroscience have illuminated the biological foundations of these deviations, showing that purchasing behavior emerges from the interplay between emotional systems and higher-order cognitive control.¹ Two forces are especially visible in this tension: loss aversion, where the pain of losing outweighs equivalent gains, and reward anticipation, where expectation of future reward provides strong motivation independent of consumption itself.^{2,3} These mechanisms help explain why consumers overpay to avoid losses, pursue anticipated pleasure through impulsive purchases, and show predictable irrationalities that defy strict economic rationality.⁴

Loss aversion was first formalized by Kahneman and Tversky through prospect theory, which proposes that losses are weighted more heavily than equivalent gains in the value function that governs risky choice.¹⁴ The behavioral parameter λ (lambda), which captures this asymmetry, is frequently estimated near 2.0, indicating that losses carry roughly twice the subjective weight of equivalent gains.² Subsequent research has proposed links between loss aversion and norepinephrine signaling, and between reward-related risk taking and dopamine.² Neuroimaging studies have shown that framing effects during decision-making are associated with amygdala activation, while reduced susceptibility to framing is associated with prefrontal cortex activity.⁵ Meanwhile, reward anticipation is associated with activity in the mesolimbic dopamine pathway, including the ventral tegmental area, nucleus accumbens, and prefrontal cortex.⁴ This

anticipatory signaling is thought to help explain why consumers browse extensively, add items to carts, and experience excitement that may exceed the value of the eventual purchase.³

Consumer neuroscience has mapped these patterns with increasing precision, but practical application remains restricted.⁶ EEG and fMRI studies have achieved 80–90% accuracy in classifying purchase intent from neural signals, but high costs, small samples, and ethical concerns limit scalability.⁷ The field identifies AI integration as a top research priority, but no systematic evaluation exists comparing neuroscience-derived decision criteria against scalable computational alternatives.⁶ This creates a substantial analytical gap: while researchers have documented the behavioral and biological basis of consumer irrationality, cost-effective tools to simulate and predict it at population scale remain limited.

Recent advances in large language models (LLMs) offer a potential solution through digital twins: personalized LLM agents calibrated on individual behavioral data. One recent study reported 86% future-purchase prediction accuracy using individualized fine-tuning and retrieval-augmented generation on Amazon consumer data.⁸ Techniques like silicon sampling use the biases absorbed during LLM training to simulate consumer panels at low cost.⁹ Yet fundamental questions remain. Can these synthetic agents replicate not just average behavioral outcomes, but the specific decision patterns associated with loss aversion and reward anticipation? Where do they systematically deviate from human decision-making?

This paper addresses that gap through a narrative review. We first establish behavioral and neurobiological benchmarks for loss aversion and reward anticipation using the consumer neuroscience literature. We then assess whether current LLM architectures reproduce the approximately 2:1 loss-gain asymmetry observed in prospect-theory studies, the temporal dynamics of reward anticipation, and the individual variation documented in behavioral and neuroimaging research. Our analysis reveals both capabilities and limitations, providing a framework for evaluating AI-based consumer models.

Contribution Statement: While recent studies have begun to examine prospect-theory-consistent behavior in LLMs, prior work has not systematically applied neuroscience-derived evaluation criteria — specifically behavioral, computational, and neurobiological benchmarks — to LLM-based consumer digital twins of the kind designed to simulate individualized purchasing behavior. This paper synthesizes consumer neuroscience and generative AI literatures to

establish explicit evaluation criteria and uses those criteria to identify where LLM consumer simulations succeed at behavioral prediction and where they have not yet demonstrated mechanistic correspondence with documented neural processes. The result is a framework that researchers can use to distinguish behavioral prediction from reproduction of underlying decision mechanisms in AI-based consumer research.

The remainder of this paper proceeds as follows. Section 2 describes the review methodology. Section 3 details the neuroscience of loss aversion, reward anticipation, and the theoretical predictions these mechanisms generate. Section 4 examines LLM-based consumer simulation, including silicon consumers, generative AI prediction, and benchmarking methods. Section 5 synthesizes findings, evaluates LLM evidence against each proposed benchmark, and discusses limitations and future research. Section 6 concludes.

2. Methods

Review Type: Narrative Review

This paper follows a narrative review design. Literature was identified through searches of Google Scholar, PubMed, IEEE Xplore, and SSRN conducted between April and June 2026 (final search: June 28, 2026). Search terms included combinations of: "loss aversion," "reward anticipation," "consumer neuroscience," "neuromarketing," "LLM digital twins," "silicon sampling," "generative AI consumer behavior," and "prospect theory lambda." No formal systematic search protocol was registered in advance.

Articles were included if they (a) reported empirical, review, or theoretical findings on loss aversion or reward anticipation in consumer or economic decision-making contexts, (b) described LLM-based methods for simulating consumer behavior, or (c) proposed benchmarking frameworks for evaluating AI behavioral simulation against human data. Sources were prioritized if they appeared in peer-reviewed journals or established conference proceedings. Preprints hosted on SSRN or arXiv and working papers were included when peer-reviewed alternatives on the same topic were unavailable, but are identified as such in the reference list and are treated with appropriate caution in the synthesis. Blog posts, news articles, and non-scholarly web sources were excluded. Duplicates identified across databases were removed manually.

A total of 22 sources met these criteria and form the evidence base for this review. Of these, 16 are peer-reviewed journal articles or conference papers, two are working papers (one SSRN/MSI, one Harvard Business School), two are published in open-access journals, and two are preprints (one arXiv, one identified through forward citation tracing). Sources were initially identified by keyword search; a second pass of forward and backward citation tracing was conducted to identify primary neuroimaging studies and directly relevant LLM behavioral studies that did not surface in initial keyword results. Source classification: peer-reviewed (16), working papers (2), open-access (2), preprints (2). Because this is a narrative rather than systematic review, the search and selection process was not exhaustive, and the findings should be interpreted as a conceptual synthesis rather than a comprehensive summary of all available evidence.

3. Neuroscience of Consumer Decision-Making

3.1 Loss Aversion

Loss aversion is one of the most well-documented biases in consumer behavior. Numerous studies show that losing money, time, or status produces stronger subjective responses than gaining the same amount. In prospect-theory studies, the behavioral loss-aversion parameter λ is frequently estimated near 2.0, indicating that the subjective weight of losses is approximately twice that of equivalent gains.² This behavioral asymmetry helps explain why consumers may avoid switching brands, overpay to keep a current subscription, or hesitate to sell products they already own, even when doing so would be economically rational.

Research in consumer neuroscience has linked this behavioral pattern to underlying neural systems. Sokol-Hessner and Rutledge proposed a framework in which norepinephrine is associated with loss aversion and the expectation of negative outcomes, while dopamine is more strongly associated with reward pursuit and risk taking for gains.² This suggests that loss aversion may emerge from the interaction between distinguishable neurochemical systems, although the precise mechanistic relationship between neurotransmitter activity and behavioral λ remains an active area of investigation.

Neuroimaging studies have provided additional evidence. De Martino and colleagues found that framing effects during risky choice were associated with amygdala activation, while individuals who were less susceptible to framing showed greater activity in orbital and medial prefrontal

cortex.⁵ This finding is consistent with dual-process accounts in which rapid emotional responses interact with slower deliberative processing, but the relationship is more nuanced than a simple "emotional amygdala versus rational prefrontal cortex" model. Multiple brain regions contribute to valuation and choice, and their interactions are context-dependent.⁵ A complementary neuroimaging study by Tom and colleagues found that a broad network of regions — including midbrain dopaminergic areas and the ventral striatum — showed increasing activity as potential gains rose but decreasing activity as potential losses rose, and that individual differences in this neural asymmetry predicted behavioral loss aversion.¹⁵ These results locate loss aversion not in a distinct fear circuit but in the same reward-valuation system that codes for gains, operating with a steeper response to potential losses.

In a marketing or shopping context, loss aversion manifests in concrete ways. Limited-time offers such as "only 2 hours left" make the potential loss of a discount feel more intense than the possible gain of saving extra money. Subscription models leverage loss aversion by making cancellation feel like losing a benefit rather than avoiding a cost. The endowment effect, where people value an object more simply because they own it, is another behavioral expression of loss aversion.¹⁰ Neuromarketing experiments using EEG, fMRI, or skin-conductance measures have found stronger physiological responses to loss-framed messages compared to gain-framed messages.⁶

Loss aversion is not uniform across individuals. Personality, risk tolerance, and clinical status all influence the magnitude of the effect. Some studies suggest that higher anxiety is associated with stronger loss-related responses, while more impulsive or risk-seeking individuals may show attenuated loss aversion.² This variation means that a successful model of consumer behavior must account for more than average behavioral tendencies and should preserve meaningful individual differences.

3.2 Reward Anticipation

If loss aversion explains why consumers avoid losses, reward anticipation explains why they pursue expected pleasure. In consumer behavior, anticipation can be as important as the outcome itself: people are frequently motivated not just by what they will receive, but by the expectation that something good is coming. The excitement of browsing, comparing, and choosing can engage reward-related neural systems before any actual product is bought or used.

At the biological level, reward anticipation is associated with dopamine activity in the mesolimbic reward pathway. This circuit includes the ventral tegmental area (VTA), the nucleus accumbens (NAcc), and the prefrontal cortex (PFC), which together support the detection, valuation, and pursuit of rewarding outcomes. Several studies have observed strong striatal activity during reward anticipation, with anticipation and receipt showing partly distinct spatial and temporal neural signatures.^{4,16} Knutson and colleagues demonstrated using event-related fMRI that anticipation of increasing monetary rewards selectively recruits the nucleus accumbens, while anticipation of punishments did not produce the same activation — a temporal dissociation that has been replicated in studies of purchase decisions and financial risk-taking.^{16,20} This helps explain why shopping can feel exciting before purchase and why that excitement may diminish after the item is obtained.³

This anticipatory process also explains many common consumer behaviors. People add items to their cart and leave them there, revisit product pages, or delay checkout while still feeling psychologically engaged with the possibility of buying. In these cases, the consumer is not making a purely utilitarian calculation but is responding to motivational signals generated by expectation. Marketing strategies take advantage of this by using scarcity, countdowns, exclusivity, and aspirational imagery, all of which may intensify the feeling that a reward is close at hand.

Reward anticipation does not operate in isolation. It interacts with systems that evaluate price, risk, identity, and social meaning. A product that signals status may trigger stronger anticipatory responses than a product with the same practical value but less social appeal.⁴ This is one reason why standard economic models tend to underestimate the role of desire, excitement, and expectation in purchase decisions. Human choice is driven not only by what people receive, but by how strongly they anticipate receiving it.

3.3 Evaluation Criteria for LLM Consumer Simulation

The neuroscience and behavioral economics literatures generate specific, testable predictions about consumer decision-making. These can be organized into two categories: behavioral benchmarks, which describe observable choice patterns, and neurobiological benchmarks, which describe the underlying physiological processes.

Behavioral benchmarks. First, the prospect-theory value function predicts a consistent asymmetry in which losses carry approximately twice the subjective weight of equivalent gains ($\lambda \approx 2.0$).² A consumer model should reproduce this asymmetry in choice data. Second, consumers should show measurable changes in behavior during anticipation phases, not only at the point of outcome. Third, individual differences in both loss sensitivity and reward responsiveness should be preserved across different consumer profiles, rather than being averaged away.

Neurobiological benchmarks. At the neural level, loss aversion has been associated with norepinephrine-mediated threat processing and amygdala activity during framing tasks, while reward anticipation has been associated with dopaminergic activity in the mesolimbic pathway.^{2,5} These associations suggest that a model claiming to simulate the mechanisms of consumer choice, rather than merely predicting its outcomes, would need to demonstrate correspondence with the temporal, spatial, or pharmacological signatures observed in neuroimaging and neurochemical studies.^{15,16} Behavioral agreement alone cannot establish equivalence of underlying mechanisms.

These two levels produce distinct evaluation criteria for LLM-based digital twins. At the behavioral level: (a) do LLMs reproduce the approximate loss-gain asymmetry? (b) do they capture anticipatory dynamics? (c) do they reflect individual differences? At the neurobiological level: do LLMs produce outputs that correspond to the physiological patterns observed in human neuroimaging studies? The distinction matters because a model that reproduces behavioral choice patterns without correspondence to underlying mechanisms is a useful predictor but not a mechanistic simulator.

4. LLM-Based Consumer Simulation

4.1 Silicon Consumers

Neuromarketing originally attempted to study consumer decisions by measuring neural activity directly.^{1,7} Studies using fMRI, EEG, eye tracking, and skin conductance showed that consumers do not always make decisions in fully rational ways; prices, brands, and advertisements can activate reward, threat, and attention systems that influence what people buy.^{7,5} Over time, machine learning models applied to these signals improved the ability to classify consumer

responses, with some reviews reporting accuracies in the 80–90% range for preference prediction.⁷ But even with these advances, neural measurement remains expensive, narrow, and difficult to scale, which has pushed researchers toward alternative approaches.⁶

This logistical bottleneck has led to the emergence of silicon sampling, in which researchers use LLMs to generate synthetic consumer responses.⁹ These models can simulate consumers exhibiting patterns consistent with bounded rationality: approximate human choices combined with systematic errors that resemble human biases.^{9,21} By creating digital twins, personalized LLM agents adapted to individual behavioral data through fine-tuning and retrieval-augmented generation, researchers can predict purchasing decisions with high accuracy in specific settings. Li, Wei, and Wang reported 86% average future-purchase prediction accuracy across 304 individualized digital twins built on Amazon consumer data.⁸ While traditional neuromarketing identifies the biological correlates of consumer behavior, LLMs provide a scalable way to predict it at a population level.^{6,8}

However, silicon consumers are not merely a cheaper substitute for traditional research. They also provide a way to test where AI can capture the decision patterns that define actual consumer choice. A useful LLM consumer should not only predict a product choice but also show sensitivity to price framing, brand loyalty, and trade-offs between utility and emotion.⁹ If a model can reproduce these behavioral patterns consistently, it can serve as a meaningful proxy for human respondents in early-stage marketing research.

At the same time, the limits of silicon consumers must be acknowledged. Current LLM-based consumer models are generally not trained to reproduce measured neural activity, and behavioral calibration does not by itself provide evidence of neurobiological correspondence.¹¹ They can reproduce consumer language and choice logic, but behavioral agreement alone cannot establish that the model is generating those outputs through the same processes as a human brain.²² Their strength lies in behavioral approximation rather than mechanistic simulation.

4.2 Generative AI Prediction

Generative AI prediction refers to the use of large language models and other generative systems to forecast what consumers are likely to do based on past behavior, preferences, and context. Compared with traditional methods, which often rely on surveys or limited statistical models,

generative AI can process much larger and more complex datasets and identify patterns that manual analysis may miss.¹²

Technically, general-purpose LLMs are predominantly pretrained on text data; consumer digital twins are then adapted using consumer-specific information through fine-tuning and retrieval-augmented generation.⁸ In marketing, this can help companies predict consumer responses before a campaign or product launch.⁹ Brand, Israeli, and Ngwe showed that LLM-based market research can generate consumer-like responses and realistic willingness-to-pay estimates, suggesting that generative AI can serve as a practical tool for faster and cheaper early-stage testing.⁹

Generative AI is also useful for personalization. These models can create marketing content tailored to a user's likely preferences.¹³ This matters because consumers do not all respond in the same way, and better predictions lead to more relevant messages and stronger engagement.

However, generative AI does not process information through the same biological systems as a human brain. It can capture statistical regularities in consumer behavior, but current LLM-based consumer models lack the biological substrate measured in human neuroimaging and neurochemical studies.¹¹ Li et al. showed that digital twins are most effective when carefully calibrated to individual data, suggesting these models are useful tools but not complete replacements for human behavioral research.⁸

The central question is whether generative AI can capture the specific decision patterns associated with loss aversion and reward anticipation, or whether it can only approximate final behavioral outcomes.⁸ Because current consumer digital-twin systems are not trained to reproduce measured neural activity, behavioral calibration alone does not provide evidence of neurobiological correspondence.

4.3 Benchmarking AI against Behavioral and Neural Evidence

To determine whether generative AI can function as a consumer model, it is not sufficient to evaluate only final purchase prediction. The real test is whether these models can match the specific behavioral and neural patterns that shape human choice. Benchmarking allows researchers to compare AI outputs with what neuroscience and behavioral economics have

documented in humans, rather than simply grading the AI on whether it predicted the correct final outcome.¹¹

At the behavioral level, researchers can use the lambda coefficient from prospect theory to measure whether an AI shows loss-gain asymmetry comparable to what is observed in human choice data.² They can also use sequential decision tasks to assess whether AI models capture temporal changes in motivation across anticipation and outcome phases, or whether they merely predict final choices.⁸

Hu et al. provided the most comprehensive benchmark to date. Their SimBench study aggregated 20 diverse behavioral datasets and found that even the best-performing LLMs achieved only modest simulation fidelity (40.80 out of 100), with particular difficulty simulating specific demographic groups.¹¹ This finding is important because it suggests that LLMs struggle most precisely where individual and group differences matter most, which is one of the three evaluation criteria proposed in Section 3.3.

Li et al. showed that individualized digital twins can predict specific shopping outcomes with high accuracy in their particular framework, but high prediction accuracy does not by itself establish that the model captures the reasons those choices were made.⁸ An AI can correctly predict what someone will buy while generating that prediction through entirely different computational processes than those documented in human neuroimaging studies. This is why separate behavioral and neurobiological benchmarks are needed: they move the evaluation beyond final outcomes and ask whether the AI reproduces the specific patterns associated with human decision-making.

Evidence on whether LLMs reproduce prospect-theory-consistent framing and loss-aversion effects at the behavioral level is mixed. Suri and colleagues experimentally demonstrated that GPT-3.5 exhibits several human-like decision heuristics — including framing effects, the endowment effect, anchoring, and representativeness — in patterns broadly consistent with the processes that underlie loss aversion and effort reduction in humans.¹⁷ However, more recent work comparing LLM predictions with human choices across 51 scenarios and 9,600 model responses against 2,104 human participants found that while GPT-3.5 matched average human risk preferences, it showed reversed framing effects — treating gain-framed options as cues for increased risk rather than caution, opposite to the pattern predicted by prospect theory — and

GPT-4 was indiscriminately risk-averse across social contexts.¹⁸ An arXiv preprint by Payne similarly found that prospect theory often anticipates LLM behavior across risky scenarios, but that context — particularly the framing of scenarios — substantially shifts the results, with military versus civilian framings producing markedly different apparent loss-aversion effects.¹⁹ Taken together, these findings suggest that LLMs can produce behavioral outputs that superficially resemble human decision patterns under risk, but that the underlying processes — and therefore the reliability of the correspondence — remain uncertain and context-dependent. This is precisely what makes the behavioral-versus-neurobiological distinction proposed in Section 3.3 essential: surface-level resemblance in behavioral outputs does not establish mechanistic correspondence.

5. Discussion

5.1 Evidence Synthesis

This review examined whether LLM-based digital twins can approximate human loss aversion and reward anticipation in consumer decision-making. Table 1 summarizes the current state of evidence for each evaluation criterion proposed in Section 3.3.

Table 1. Summary of evaluation criteria: human evidence versus LLM evidence.

Criterion	Human Evidence	LLM Evidence	Status	Key Sources
Loss/gain asymmetry ($\lambda \approx 2.0$)	Well-established in behavioral prospect-theory studies	No direct λ benchmark reported for consumer LLM twins	Not established	Sokol-Hessner & Rutledge (2019)
Anticipatory temporal dynamics	Neuroimaging shows distinct anticipation vs. receipt responses in striatal regions	Minimal direct temporal comparison in LLM studies	Not established	Peterson (2005); Evans (2026)
Individual variation	Significant person- and context-dependent variability in λ and reward sensitivity	Personalized twins show some behavioral personalization; SimBench finds difficulty with demographic subgroups	Partially supported	Li et al. (2025); Hu et al. (2026)
Purchase prediction accuracy	Human transaction outcomes as ground truth	86% in one individualized consumer twin study (Amazon data, 304 twins)	Supported in specific setting	Li et al. (2025)
Overall behavioral simulation fidelity	Human behavioral variability across 20 diverse datasets	Best LLM achieved 40.80/100; difficulty with specific demographic	Limited	Hu et al. (2026)

		groups		
--	--	--------	--	--

As Table 1 shows, the strongest current evidence for LLM consumer digital twins lies in purchase-prediction accuracy within specific, well-calibrated settings. The weakest areas are those most relevant to this review's central question: whether LLMs reproduce the behavioral asymmetries and temporal dynamics associated with loss aversion and reward anticipation. Recent studies have begun to examine LLM behavior under loss-framing and risk paradigms, with mixed findings: some work finds framing effects consistent with human patterns, while others find reversed or context-dependent effects.^{17,18,19} However, these studies have not examined individualized consumer digital twins of the kind built by Li et al., and no study to date has directly measured whether such a consumer twin produces a loss-aversion coefficient comparable to human prospect-theory estimates or exhibits the temporal dynamics of reward anticipation documented in neuroimaging literature.

The gap between prediction and mechanistic correspondence is where current generative AI most clearly falls short. A model can generate accurate behavioral predictions while producing those predictions through entirely different computational processes than those documented in human neuroimaging studies. Metrics such as the lambda coefficient and temporal response profiling could help researchers determine where LLM outputs align with human decision patterns at a deeper level than final outcomes.^{2,11}

5.2 Strengths and Limitations

This review has several strengths. It bridges two literatures, consumer neuroscience and generative AI, that are rarely connected in existing work. It distinguishes behavioral from neurobiological benchmarks and proposes explicit evaluation criteria. It also provides a synthesis table that future researchers can extend as new evidence becomes available.

However, several limitations should be acknowledged. First, this is a narrative review with 22 included sources, which is broader than the original submission but still does not constitute a comprehensive coverage of any single literature it draws on. The literature search was not registered in advance, the search strategy used keyword combinations and citation tracing rather than formal Boolean strings with pre-specified databases, and the selection of sources was not

exhaustive. Important work may have been missed, and the synthesis reflects the judgment of the authors rather than a reproducible screening protocol. The neuroscience sections in particular rely primarily on a small set of landmark studies — the primary neuroimaging literature on loss aversion and reward anticipation is substantially larger than what is reviewed here. Future work should engage more comprehensively with meta-analyses of loss-aversion neural correlates and the growing experimental literature on LLM prospect-theory behavior.

Second, this paper is entirely theoretical: no original experiments, simulations, or quantitative analyses were conducted. The evaluation criteria proposed here have not been empirically tested, and the conclusions about LLM limitations are inferred from existing studies rather than from direct experimentation. Third, the literature on LLM-based consumer simulation is very recent. Several of the sources reviewed are preprints or working papers that have not yet undergone extensive independent replication. The Li et al. digital-twin result (86% accuracy) comes from one specific framework with one dataset, and the SimBench fidelity score (40.80/100) aggregates across diverse tasks that may not all be directly relevant to consumer behavior. Both findings should be treated as preliminary. Fourth, this review does not address all cognitive biases relevant to consumer behavior. Loss aversion and reward anticipation were selected as focal mechanisms, but anchoring, social proof, and other biases also influence purchasing decisions.

5.3 Implications and Future Research

The framework presented here has implications for both researchers and practitioners. For researchers, the separation of behavioral from neurobiological benchmarks provides a starting point for designing experiments that go beyond accuracy metrics. Future work should directly test whether LLMs produce lambda coefficients comparable to those observed in human prospect-theory experiments, whether they exhibit distinct response patterns across anticipation and outcome phases in sequential purchasing scenarios, and whether they can simulate meaningful individual differences when calibrated on diverse consumer profiles.

For marketing practitioners, the findings suggest that LLM-based consumer simulations are best suited for early-stage research where speed and scale matter more than mechanistic understanding. They should not be treated as replacements for studies that require understanding of emotional or neurobiological drivers. When the goal is to understand why consumers make

certain choices rather than predicting what they will choose, neuroscience-based methods remain necessary.

6. Conclusion

This review argues that LLM-based digital twins can approximate broad behavioral patterns of consumer choice, but current evidence does not establish that they reproduce the mechanisms associated with loss aversion and reward anticipation. At the behavioral level, prospect-theory research documents a consistent asymmetry in which losses carry roughly twice the subjective weight of gains ($\lambda \approx 2.0$).² At the neurobiological level, loss aversion has been linked to norepinephrine-mediated threat processing and amygdala involvement, while reward anticipation has been linked to dopaminergic activity in the mesolimbic pathway.^{2,4,5} Both mechanisms vary across individuals and operate through biological processes that current LLM architectures are not designed to model.

One consumer digital-twin study has demonstrated strong prediction accuracy in a specific setting, and a large-scale behavioral benchmark has found that LLM simulation fidelity remains modest overall. Neither finding addresses whether LLMs reproduce the specific behavioral asymmetries or temporal dynamics documented in consumer neuroscience. Until future research directly tests the evaluation criteria proposed here, LLM-based consumer models should be treated as useful behavioral predictors rather than mechanistic simulators of human decision-making.

Abbreviations

AI: Artificial intelligence; EEG: Electroencephalography; fMRI: Functional magnetic resonance imaging; LLM: Large language model; NAcc: Nucleus accumbens; PFC: Prefrontal cortex; RAG: Retrieval-augmented generation; VTA: Ventral tegmental area; λ : Lambda (prospect-theory loss-aversion parameter)

Declarations

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethics Statement

Ethics approval was not required. This study is a narrative literature review and did not involve human participants, private identifiable data, animal subjects, or primary data collection.

Data Availability Statement

No original data were generated or analyzed in this study. All sources reviewed are cited in the reference list and are publicly accessible through the databases and repositories identified in Section 2.

AI Use Statement

The authors used Claude (Anthropic) as a writing and brainstorming aid during manuscript preparation, including assistance with outlining, drafting, editing, formatting, and reference verification. All AI-generated content was reviewed, revised, and verified by the authors, who take full responsibility for the accuracy, originality, and integrity of the manuscript. Generative AI was not listed as an author.

Author Contributions

P. Nagarajan contributed to conceptualization, literature review, methodology, and writing of the original draft. S. Tibrewal contributed to conceptualization, literature review, critical revision, and editing. Both authors reviewed and approved the final manuscript.

References

1. Hubert M, Kenning P. A current overview of consumer neuroscience. *J Consum Behav*. 2008;7(4–5):272–292. doi:10.1002/cb.251

2. Sokol-Hessner P, Rutledge RB. The psychological and neural basis of loss aversion. *Curr Dir Psychol Sci.* 2019;28(1):20–27. doi:10.1177/0963721418806510
3. Evans RD Jr. Dopamine, the master reward neurotransmitter: research propositions for advancing consumer neuroscience and neuromarketing research. *J Contemp Mark Sci.* 2026;9(1):49–62. doi:10.1108/JCMARS-12-2024-0049
4. Peterson RL. The neuroscience of investing: fMRI of the reward system. *Brain Res Bull.* 2005;67(5):391–397. doi:10.1016/j.brainresbull.2005.06.015
5. De Martino B, Kumaran D, Seymour B, Dolan RJ. Frames, biases, and rational decision-making in the human brain. *Science.* 2006;313(5787):684–687. doi:10.1126/science.1128356
6. Gupta R, Kapoor AP, Verma HV. Neuro-insights: a systematic review of neuromarketing perspectives across consumer buying stages. *Front Neuroergon.* 2025;6:1542847. doi:10.3389/fnrgo.2025.1542847
7. Rawnaque FS, Rahman KM, Anwar SF, et al. Technological advancements and opportunities in neuromarketing: a systematic review. *Brain Inform.* 2020;7:10. doi:10.1186/s40708-020-00109-x
8. Li B, Wei Q, Wang X. Predicting behaviors with large language model (LLM)-powered digital twins of consumers. MSI Working Paper No. 25-135. 2025.
9. Brand J, Israeli A, Ngwe D. Using LLMs for market research. Harvard Business School Working Paper No. 23-062. 2024.
10. Rolls ET. Emotion and reasoning in human decision-making. *Economics.* 2019;13(2019–39):1–31. doi:10.5018/economics-ejournal.ja.2019-39
11. Hu T, Baumann J, Lupo L, Collier N, Hovy D, Röttger P. SimBench: benchmarking the ability of large language models to simulate human behaviors. In: *International Conference on Learning Representations (ICLR)*; 2026. arXiv:2510.17516.
12. Madanchian M. Generative AI for consumer behavior prediction: techniques and applications. *Sustainability.* 2024;16(22):9963. doi:10.3390/su16229963
13. Lee GH, Lee KJ, Jeong B, Kim TK. Developing personalized marketing service using generative AI. *IEEE Access.* 2024;12:22394–22402. doi:10.1109/ACCESS.2024.10419357

14. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica*. 1979;47(2):263–291. doi:10.2307/1914185
15. Tom SM, Fox CR, Trepel C, Poldrack RA. The neural basis of loss aversion in decision-making under risk. *Science*. 2007;315(5811):515–518. doi:10.1126/science.1134239
16. Knutson B, Adams CM, Fong GW, Hommer D. Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *J Neurosci*. 2001;21(16):RC159. doi:10.1523/JNEUROSCI.21-16-j0002.2001
17. Suri G, Slater LR, Ziaee A, Nguyen M. Do large language models show decision heuristics similar to humans? A case study using GPT-3.5. *J Exp Psychol Gen*. 2024;153(4):1066–1075. doi:10.1037/xge0001547
18. Phelps M, Capraro V, Holt N, Savulescu J, Shanahan M. Evaluating the ability of large language models to predict human social decisions. *Sci Rep*. 2025;15:21428. doi:10.1038/s41598-025-17188-7
19. Payne K. An analysis of AI decision under risk: prospect theory emerges in large language models. *arXiv preprint arXiv:2508.00902*. 2025.
20. Knutson B, Rick S, Wimmer GE, Prelec D, Loewenstein G. Neural predictors of purchases. *Neuron*. 2007;53(1):147–156. doi:10.1016/j.neuron.2006.11.010
21. Horton JJ. Large language models as simulated economic agents: what can we learn from homo silicus? NBER Working Paper No. 31122. 2023. doi:10.3386/w31122
22. Brady G, Daly A, Vallgård A, Frith CD. Dual-process theory and decision-making in large language models. *Nat Rev Psychol*. 2025;4:248–260. doi:10.1038/s44159-025-00506-1