

When AI Starts Spending: Examining the Impact of Agentic Artificial Intelligence on Consumer Financial Decision-Making

-
Tvisha Tuli

Abstract

The emergence of agentic artificial intelligence (AI) is transforming online shopping by enabling AI systems to search, compare, select, and purchase products on behalf of consumers. This paper examines how delegating purchasing decisions to AI agents influences consumer financial behaviour, spending awareness, decision-making autonomy, and accountability. Using an extended interdisciplinary literature review, it synthesizes research from behavioural economics, consumer psychology, cybersecurity, and consumer protection law. The findings highlight potential benefits, including improved price comparisons, reduced search costs, and automated budget management. However, significant risks include reduced spending awareness, automation bias, algorithmic manipulation, privacy vulnerabilities, and unclear liability for mistaken purchases. The review also identifies concerns regarding young consumers and the development of financial literacy. While existing research establishes links between less visible payment mechanisms and increased spending, direct empirical evidence on AI-driven purchasing remains limited. The paper recommends transparent recommendations, spending limits, meaningful human approval, and stronger consumer protection regulations.

Keywords: Agentic AI, Agentic Commerce, Consumer Financial Behaviour, Spending Awareness, Behavioural Economics, Financial Literacy, Consumer Protection.

1. Introduction

Artificial intelligence is moving from answering shopping questions to making purchases itself. For most of the past decade, an AI system could tell a consumer which blender had the best reviews, but a person still had to click “buy,” type a card number, and confirm the order. In 2025 and 2026, that final step began to disappear. This paper asks what the shift from AI that advises to AI that acts does to the way people decide to spend money.

1.1 A Purchase Nobody Saw

Consider a student who types a single sentence into an AI assistant: “Order me running shoes under \$80 that arrive by Friday.” Within minutes, the assistant has searched several stores, chosen a pair, applied a saved card, and confirmed the order. The shoes meet every stated requirement. Yet the student never saw the other options, never compared colors or brands, and never paused at a checkout page to ask whether new shoes were needed at all.

The scenario raises the question at the center of this paper: if you never see the choices, did you really make the decision? Economists have traditionally treated the act of choosing as the moment when a consumer’s preferences are expressed and tested. When an AI agent performs the search, the comparison,

and the payment, the human contribution shrinks to a sentence of instructions. The purchase may be efficient, but it is no longer obvious whose judgment it reflects.

1.2 Background: From Chatbots to Agents

A chatbot answers questions; an agentic AI system takes actions. Researchers at OpenAI describe “agenticness” as the degree to which a system can pursue complex goals in complex environments with limited direct supervision (Shavit et al. 2023). In commerce, this means software that can open websites, fill a shopping cart, enter shipping details, and pay.

Several products launched in 2025 turned that idea into consumer features. OpenAI released Operator, a browser-using agent, as a research preview in January 2025 and folded its capabilities into ChatGPT agent in July 2025 (OpenAI 2025a; OpenAI 2025b). In September 2025, OpenAI introduced Instant Checkout, which lets users buy items from participating merchants without leaving a ChatGPT conversation, built on an open Agentic Commerce Protocol developed with the payment company Stripe (OpenAI 2025c). Amazon began testing “Buy for Me,” a feature in which an AI agent completes purchases on other brands’ websites for a customer, in April 2025 (Amazon 2025).

Payment networks moved at the same time. Mastercard announced Agent Pay and Visa announced Visa Intelligent Commerce on consecutive days in late April 2025, each promising tokenized credentials that let AI agents pay without seeing a customer’s real card number (Mastercard 2025; Visa 2025). In September 2025, Google released the Agent Payments Protocol (AP2), an open standard intended to give merchants and banks verifiable proof that a user authorized an agent’s purchase (Google Cloud 2025).

The timing matters. When payment companies build the infrastructure for agents to pay safely, AI purchasing moves from experiment toward everyday feature. Consultants at McKinsey have estimated that agentic commerce could orchestrate as much as \$1 trillion in US retail revenue by 2030 (McKinsey & Company 2025). Forecasts from firms that advise the industry should be treated with caution, but they show how seriously businesses are taking the shift.

1.3 Working Thesis

This paper argues that agentic AI can make spending faster and, in some cases, cheaper, but that it removes the pauses in which people compare options and reconsider purchases. Those pauses are not simply inefficiencies. Behavioral research suggests they help people notice what they spend and catch decisions they would later regret. Whether agentic commerce helps or harms consumers therefore depends less on how intelligent the agent is than on the safeguards built around it: spending limits, transparency about how products are chosen, and meaningful human approval before money moves.

1.4 Approach: An Extended Literature Review

Evidence on how AI agents affect spending is scattered across consumer psychology, human factors research, computer security, law, and a fast-growing body of company announcements and news reports. Rather than collect new data, this paper is an extended literature review: it gathers research from these fields, evaluates its quality, and integrates it into a single argument about how agentic AI changes consumer decisions (Torraco 2005; Snyder 2019). Unlike a systematic review, which follows a fixed protocol to count and screen every study on a narrow question, an extended review is designed for broad, interdisciplinary topics where the goal is to build understanding rather than to tally results (Grant and Booth 2009). Because much of the evidence on agentic commerce appears outside academic journals, the review also draws on carefully evaluated grey literature from regulators, companies, and security researchers (Garousi, Felderer, and Mäntylä 2019).

1.5 Roadmap

Section 2 sets out the review questions and propositions, and Section 3 defines key terms. Section 4 summarizes the background theory and presents the conceptual framework that organizes the review. Section 5 explains the review approach: how sources were found, selected, evaluated, and synthesized. Section 6 reports the results theme by theme, covering the evidence base, the benefits and risks of agents, conflicts of interest, legal responsibility, and effects on young people. Section 7 answers the review questions and identifies gaps in the evidence. Section 8 addresses ethical considerations, limitations, and counterarguments, and Section 9 concludes with recommendations.

2. Review Questions and Propositions

The central review question is: *What does existing research reveal about how handing purchases to agentic AI changes the quality, awareness, and control of consumers' financial decisions?* In this paper, quality refers to whether a purchase matches the person's real needs and budget; awareness refers to whether the person knows what was bought and what it cost; and control refers to whether the person can direct, limit, and reverse the agent's actions.

2.1 Sub-questions

Five sub-questions break the main question into parts that the literature can answer:

1. Which purchases are people willing to let an AI make, and does the price or type of item change that willingness?
2. Do people who use AI agents spend more, or keep less accurate track of what they spend?
3. Whose interests does the agent serve: the user, the seller, or the company that built it?
4. Who is responsible when an agent buys the wrong thing or is tricked by a scam?
5. Are teenagers and young adults affected differently from older consumers?

2.2 Propositions

A literature review does not test hypotheses with new data, but stating expectations in advance makes the synthesis more disciplined: the evidence is checked against them rather than gathered to fit them. Two propositions guide the review.

Proposition 1 (P1): The literature will indicate that people readily delegate cheap, routine purchases, such as groceries and household refills, but resist delegating expensive or personal purchases, such as electronics, gifts, and clothing. This expectation draws on research showing that reliance on automation depends on how risky an error would be (Lee and See 2004). A mistaken order of paper towels costs little; a mistaken laptop or an ill-chosen gift costs money and carries social meaning.

Proposition 2 (P2): The literature will indicate that delegating purchases weakens people's awareness and memory of how much they spend. This expectation builds on Dilip Soman's finding that people who paid by credit card, and therefore never wrote down the amount, remembered their recent expenses less accurately than people who paid by check (Soman 2001). An agent removes even the act of looking at a checkout total.

Section 7.2 evaluates both propositions against the evidence.

2.3 Review Scope

Four elements define the scope of the review: the population studied, the phenomenon of interest, the context, and the outcomes that matter for the review question.

- **Population:** consumers who shop online, including teenagers and young adults.

- **Phenomenon of interest:** handing purchase decisions or payments to automated systems, especially AI agents, along with closely related precursors such as recommendation engines, one-click payment, and autopay.
- **Context:** retail and everyday consumer purchasing, mainly in the United States, with comparative evidence from the European Union.
- **Outcomes:** amount spent, awareness and recall of spending, trust and reliance, decision quality, errors and security incidents, and the allocation of legal responsibility.

3. Key Terms

The argument relies on terms from computer science, behavioral economics, law, and research methodology. Table 1 defines them in plain English so that readers from any background can follow the discussion.

Table 1. Key terms used in this paper

Term	Plain-English definition
Agentic AI	AI that pursues a goal by taking several steps on its own, such as searching, comparing, and paying, rather than just answering a question (Shavit et al. 2023).
Agentic commerce	Shopping in which an AI agent finds, chooses, or buys products for a person.
Delegation	Handing part or all of a decision to someone or something else.
Human-in-the-loop	A design in which a person must approve key steps, such as the final payment.
Friction	Small steps, like typing a card number, that slow a purchase and give people time to reconsider.
Pain of paying	The mild discomfort people feel when money leaves their hands; weaker with cards and one-click buying (Prelec and Loewenstein 1998).
Automation bias	The tendency to trust an automated system's output without checking it (Skitka, Mosier, and Burdick 1999).
Choice architecture	The way options are presented, which shapes what people pick (Thaler and Sunstein 2008).
Tokenized credentials	A stand-in card number that an agent uses so it never sees the real card details.
Prompt injection	Hidden instructions on a website or in a document that trick an AI agent into doing something the user did not want (Greshake et al. 2023).
Financial literacy	The knowledge and skills needed to manage money well (Lusardi and Mitchell 2014).
Extended literature review	A long-form review that gathers, evaluates, and synthesizes research from several fields to answer a broad question, using a transparent but flexible

	search rather than a fixed screening protocol (Grant and Booth 2009; Snyder 2019).
Integrative synthesis	Combining findings from different kinds of studies into a new framework or perspective on a topic (Torraco 2005).
Grey literature	Material published outside academic journals, such as government reports, company announcements, and security research blogs (Adams, Smart, and Huff 2017).
Citation searching (snowballing)	Finding more sources by following the reference lists of key works and checking newer works that cite them (Wohlin 2014).

4. Background and Conceptual Framework

Before turning to the review itself, this section summarizes the established theories it draws on. Research relevant to agentic commerce comes from fields that rarely speak to one another: consumer behavior, behavioral economics, human factors engineering, finance, and law. The six themes below move from how people make purchasing decisions to how AI agents change that process and what rules govern it. Together they form the conceptual framework that organizes the synthesis in Section 6.

4.1 How Consumers Make Decisions

Marketing scholars have long described buying as a five-stage process: recognizing a need, searching for information, evaluating alternatives, making the purchase, and reflecting afterward (Engel, Kollat, and Blackwell 1968; Kotler and Keller 2016). The model is a simplification, since many everyday purchases skip stages, but it is useful here because it shows exactly which stages an AI agent can take over. A shopping chatbot mainly assists with search. An agent that buys can absorb search, evaluation, and purchase, leaving the human only the first stage (stating the need) and the last (living with the result).

Real consumers do not move through these stages like perfect calculators. Herbert Simon (1955) argued that people have limited time, information, and mental capacity, so they “satisfice,” choosing an option that is good enough rather than searching for the best. This idea, known as bounded rationality, is often used to argue in favor of AI agents: if humans can compare only a handful of products, a machine that compares thousands should do better.

Research on heuristics and biases complicates that argument. Amos Tversky and Daniel Kahneman showed that people rely on mental shortcuts, such as judging how likely something is by how easily examples come to mind, and that these shortcuts produce predictable errors (Tversky and Kahneman 1974; Kahneman 2011). Such biases do not disappear when a machine enters the picture. Instead, as later sections show, they can shift from the choice itself to the decision about whether to trust the machine making the choice.

Research on the number of options adds another wrinkle. In a well-known field study, shoppers offered 24 varieties of jam were less likely to buy than shoppers offered 6 (Iyengar and Lepper 2000). A later meta-analysis, however, found that the average “choice overload” effect across many studies was close to zero, with strong effects appearing only under particular conditions (Scheibehenne, Greifeneder, and Todd 2010). Narrowing choices can relieve people, but it is not automatically beneficial.

4.2 The Psychology of Paying

A second body of research examines what happens at the moment money leaves a person’s hands. Drazen

Prelec and George Loewenstein (1998) proposed that consumers experience an immediate “pain of paying” that partly offsets the pleasure of buying. This discomfort works as a brake on spending. Brain-imaging research supports the idea: when shoppers saw prices they considered excessive, activity rose in the insula, a region associated with unpleasant feelings, and that activity helped predict decisions not to buy (Knutson et al. 2007).

Payment methods that make spending less noticeable appear to weaken the brake. Richard Feinberg (1986) found that the mere presence of credit card logos increased how much people were willing to spend. In a widely cited experiment, Prelec and Simester (2001) found that MBA students’ willingness to pay for sold-out basketball tickets was substantially higher when they were told to pay by credit card rather than cash. Soman (2001) showed that people who paid by card recalled their recent spending less accurately, and later work found that paying with cash increases people’s emotional connection to a product after purchase (Shah et al. 2016). The size of these effects varies across studies, but the general pattern, in which less visible payment means weaker restraint, is well supported.

Online retail has spent decades removing payment friction. Amazon patented one-click ordering in 1999 (US Patent No. 5,960,411), and saved cards, digital wallets, and subscriptions have since become standard. Each step made buying easier. Agentic AI extends this trend in an important way: earlier tools removed friction from paying, while agents remove friction from choosing as well.

4.3 Trust in Automation

Whether agentic commerce helps or hurts depends heavily on how much people trust it. Human factors researchers have studied trust in automation for decades, mostly in aviation and industrial control. Parasuraman and Riley (1997) distinguished among use, misuse, disuse, and abuse of automation, where “misuse” means relying on automation when it should not be trusted. One form of misuse, automation bias, is the tendency to accept an automated system’s output without seeking contradictory information. In experiments, people using an automated aid made errors both of omission, missing problems the aid did not flag, and of commission, following the aid’s incorrect advice (Skitka, Mosier, and Burdick 1999). A later review concluded that these effects appear in experts as well as novices and are not easily trained away (Parasuraman and Manzey 2010).

The opposite pattern also exists. Dietvorst, Simmons, and Massey (2015) found that people who watched an algorithm make a mistake became more likely to choose a human forecaster, even when the algorithm was more accurate overall, a pattern they called “algorithm aversion.” Other studies find “algorithm appreciation,” with ordinary people often preferring algorithmic advice to human advice (Logg, Minson, and Moore 2019). A follow-up study offered a practical bridge: people were far more willing to use an imperfect algorithm if they could adjust its output, even slightly (Dietvorst, Simmons, and Massey 2018). John Lee and Katrina See (2004) argued that the goal should be “appropriate reliance,” meaning trust that matches what a system can actually do. For agentic commerce, the lesson is that both over-trust and under-trust carry costs. Design choices, especially whether users can review and adjust an agent’s choices, influence which one occurs.

4.4 Earlier Forms of Automated Money Decisions

Agents are not the first technology to automate spending decisions, and earlier tools offer evidence about the trade-off between convenience and control. Recommendation engines are the most familiar. Amazon’s item-to-item recommendation system has shaped what shoppers see for more than two decades (Smith and Linden 2017). Early research on online “decision aids” found that tools which screened and sorted products reduced search effort and helped shoppers find better-fitting items (Häubl and Trifts 2000). A

broad review noted that the influence of recommendation agents depends on how transparent and trustworthy they appear (Xiao and Benbasat 2007).

Automatic payments and subscriptions show the other side. Defaults are powerful: when employers automatically enrolled workers in retirement savings plans, participation rose dramatically because most people kept whatever option they were given (Madrian and Shea 2001). The same inertia that helps people save can keep them paying for subscriptions they no longer use. US regulators have repeatedly targeted designs that make signing up easy and cancelling hard. In September 2025, Amazon agreed to pay \$2.5 billion to settle Federal Trade Commission (FTC) charges that it enrolled customers in Prime without clear consent and made cancellation difficult (Federal Trade Commission 2025).

Robo-advisors, which invest money automatically according to rules, offer a more encouraging case. A study of an Indian robo-advising service found that it helped poorly diversified investors spread their risk and reduced some common behavioral biases, though it offered little benefit to investors who were already diversified (D'Acunto, Prabhala, and Rossi 2019). Taken together, these examples suggest a pattern: automation tends to help when it enforces a goal the person already holds, and to harm when it makes spending easier than the person would choose on reflection.

4.5 Agentic Commerce, 2024–2026

The products described in Section 1.2 represent three distinct models of agentic commerce. The first is the general-purpose agent, such as ChatGPT agent, which operates a web browser much as a person would and can, in principle, shop on any site. The second is in-platform checkout, such as Instant Checkout, in which merchants connect directly to the AI platform through a protocol so that the agent does not need to navigate a website. The third is the retailer-run agent, such as Amazon's Buy for Me, in which a store's own AI completes purchases for the customer.

Underneath all three, payment companies are building infrastructure meant to make agent payments safer. Mastercard's Agent Pay and Visa Intelligent Commerce both rely on tokenization, in which an agent receives a substitute card number that can be restricted to particular merchants, amounts, or time periods (Mastercard 2025; Visa 2025). Google's AP2 uses cryptographically signed "mandates." An Intent Mandate records what the user asked for, such as a price ceiling, and a Cart Mandate records the specific items the user approved, creating an audit trail if a dispute arises (Google Cloud 2025). The protocol supports both "human-present" purchases, where the user approves the final cart, and "human-not-present" purchases, where the agent buys later, for example when a price drops.

Consumer interest appears real but uneven. Adobe Analytics reported that traffic from generative AI tools to US retail websites increased roughly 1,200 percent between July 2024 and February 2025, though it remained a small share of total retail traffic (Adobe 2025). At the same time, a 2025 Pew Research Center survey found that half of American adults were more concerned than excited about the growing use of AI in daily life (Pew Research Center 2025). Most claims about the benefits of agentic commerce come from the companies selling the technology. Their announcements are valuable primary sources for what products do, but they are promotional, and independent evaluations remain scarce.

4.6 Rules and Consumer Protection

US law already protects consumers against unauthorized payments. Under the Truth in Lending Act, a cardholder's liability for unauthorized credit card use is capped at \$50 (15 U.S.C. § 1643; 12 C.F.R. § 1026.12(b)). The Electronic Fund Transfer Act and its implementing rule, Regulation E, limit liability for unauthorized debit transactions to \$50 if the consumer reports the problem within two business days and \$500 if reported later, with greater exposure if the consumer fails to report within 60 days of receiving a

statement (15 U.S.C. § 1693g; 12 C.F.R. § 1005.6). The Consumer Financial Protection Bureau (CFPB) writes and enforces these payment rules, while the FTC polices unfair and deceptive business practices under Section 5 of the FTC Act (15 U.S.C. § 45).

The European Union’s Artificial Intelligence Act, which entered into force in August 2024, offers an example of broader AI regulation. It bans AI systems that use manipulative or deceptive techniques to materially distort people’s behavior, or that exploit vulnerabilities linked to age, in ways likely to cause significant harm (Regulation (EU) 2024/1689, art. 5). It also requires that people be told when they are interacting with an AI system (art. 50).

These frameworks were written for a world in which a payment was either authorized by the account holder or stolen by someone else. Agentic commerce creates a third category: a purchase the consumer authorized in general terms but did not approve in its specifics. When an agent follows instructions but buys the wrong thing, it is not clear that the purchase counts as “unauthorized.” Section 6.5 returns to this gap.

Figure 1 brings this background together as the conceptual framework for the review. Each of the five buying stages is paired with what an agent can take over at that stage and the research sub-question that stage raises.

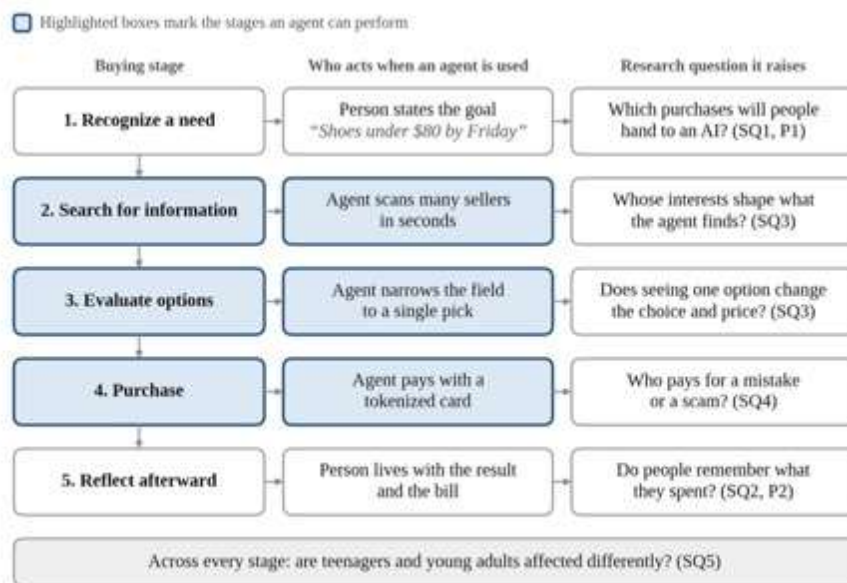


Figure 1. Conceptual framework. The person sets the goal and lives with the result, while an agent can perform the search, evaluation, and purchase in between. Each stage raises one of the sub-questions from Section 2; highlighted boxes mark the stages an agent can perform.

5. Methods: Extended Review Approach

This paper is an extended literature review rather than a study that collects new data from participants. Snyder (2019) distinguishes three broad approaches to reviewing research. Systematic reviews answer narrow questions by screening every eligible study against fixed criteria. Semi-systematic or narrative reviews trace how a broad topic has developed across research traditions. Integrative reviews critically combine different kinds of evidence to produce new frameworks or perspectives. This review combines the second and third approaches. Its topic spans psychology, economics, computer security, and law, fields with very different methods and vocabularies, and research on AI shopping agents is too new and fast-

moving for a fixed screening protocol to capture well. The aim is therefore to build a well-supported understanding of how agentic AI changes consumer decisions, not to count studies (Torraco 2005).

5.1 Finding Sources

Academic sources were located through Google Scholar and the JSTOR and EBSCO databases available through the school library, supplemented by the open-access repositories arXiv and SSRN for recent computer science and law papers not yet published in journals. Grey literature came from US regulators (the Consumer Financial Protection Bureau and the Federal Trade Commission), international bodies (the OECD), survey organizations (Pew Research Center), the newsrooms of companies building agentic commerce products (OpenAI, Google, Amazon, Visa, Mastercard, and Stripe), the blogs of security research firms, and major news outlets (Reuters, the Wall Street Journal, the Washington Post, and The Verge). Following guidance on using grey literature responsibly, these sources were relied on mainly where peer-reviewed research does not yet exist, such as descriptions of new products and documented security incidents (Adams, Smart, and Huff 2017).

Searches paired terms for the technology or behavior with terms for the outcome. Typical search terms included “agentic commerce,” “AI shopping agent,” “automation bias,” “pain of paying,” “prompt injection,” “unauthorized transaction,” and “financial literacy,” combined with words such as consumer, spending, purchase, and adolescents. Foundational studies were then traced through citation searching: following the reference lists of key works backward and checking newer works that cite them forward (Wohlin 2014). This is how older but still central research, such as studies of payment psychology from the 1980s and 1990s, entered the review.

5.2 Selecting Sources

Sources were selected for relevance and credibility rather than against a fixed checklist. Table 2 summarizes the guidelines used.

Table 2. Source selection guidelines

Guideline	Sources favored	Sources avoided
Relevance	Research on delegating purchases or payments to automated systems, payment psychology, trust in automation, consumer protection for automated transactions, and young people’s financial literacy	AI in business purchasing, automated stock trading, or other uses unrelated to consumer purchasing
Credibility	Peer-reviewed articles and books; preprints from recognized research groups; reports from regulators and international bodies; official company documentation; security research disclosures; reputable journalism	Opinion pieces without evidence; marketing copy without product details; anonymous or unverifiable sources
Currency	For agentic AI, the most recent sources available, mainly from 2023 to 2026; for established theory, the original or most influential studies	Outdated descriptions of products that have since changed

Balance	Studies that challenge the thesis as well as studies that support it	Evidence chosen only because it confirms the argument
---------	--	---

The balance guideline matters most. A common weakness of narrative reviews is that authors cite evidence supporting their argument and overlook evidence against it (Baumeister and Leary 1997). To guard against this, the review deliberately sought contrary findings, such as research on algorithm aversion and the meta-analysis questioning choice overload, and reports them alongside supporting work.

5.3 Evaluating Sources

Each source was evaluated before it was used. Four questions guided the evaluation of peer-reviewed empirical studies: Is the method clearly described? Is the sample adequate for the claim? Do the conclusions follow from the data? Are funding sources and conflicts of interest disclosed? Grey literature was evaluated with the AACODS checklist, which asks about a source's authority, accuracy, coverage, objectivity, date, and significance (Tyndall 2010). Company announcements received particular scrutiny on objectivity: they are treated as reliable evidence of what a company says its product does, but not as evidence of whether the product benefits consumers. These evaluations feed the strength-of-evidence ratings in Section 7.1.

5.4 Analyzing and Synthesizing the Literature

For each source, a notes table recorded the citation, type of source, method, main findings, limitations, and the review question it addresses. Findings were then organized under the themes of the conceptual framework in Figure 1 and compared across fields: where studies agreed, where they conflicted, and where findings from one field, such as research on credit cards, could reasonably be extended to AI agents. Following guidance for integrative reviews, the synthesis goes beyond summarizing individual studies to draw connections between them and identify what remains unknown (Torraco 2005; Snyder 2019). Because the sources use very different methods and measures, the synthesis is narrative rather than statistical.

6. Findings: Synthesis of the Literature

This section reports what the included sources show. It begins with an overview of the evidence base and then presents five themes that follow the conceptual framework: the benefits agents can offer, the risks they introduce, the conflicts of interest that shape their choices, the question of who is responsible for their mistakes, and their effects on young people.

6.1 Overview of the Evidence Base

The included evidence falls into four broad types that differ sharply in strength (Table 3). The most rigorous evidence concerns human behavior in general: decades of peer-reviewed experiments and field studies on payment psychology, defaults, and trust in automation. Evidence about AI shopping agents specifically is far thinner and newer. It consists mainly of preprints, benchmark tests, and security disclosures from 2023 to 2025, together with company announcements whose purpose is partly promotional.

Table 3. Types of evidence included in the review

Evidence type	Examples	What it contributes	Main limitation
Peer-reviewed behavioral research	Prelec and Loewenstein 1998; Soman 2001; Skitka, Mosier, and Burdick 1999; Madrian and Shea 2001	Well-tested findings on how people pay, trust automation, and respond to defaults	Predates AI agents, so findings apply by analogy
Recent AI and security research	Greshake et al. 2023; Debenedetti et al. 2024; Allouah et al. 2025; Bansal et al. 2025	Direct tests of how AI agents behave and how they can be manipulated	Often preprints or simulations rather than studies of real consumers
Legal and regulatory sources	Federal payment statutes and regulations; FTC and CFPB actions; the EU AI Act	How current law assigns responsibility and where it is silent	Describes rules, not consumer outcomes
Industry and news sources	OpenAI 2025c; Google Cloud 2025; Fowler 2025; Guardio Labs 2025	What products do and documented failures in practice	Promotional bias in company sources; anecdotal news reports

The most striking feature of the evidence base is a gap. None of the included sources is a peer-reviewed study measuring how real consumers' spending, awareness, or satisfaction changes when they use an AI agent to make purchases. The themes below therefore combine strong evidence about human behavior with weaker but direct evidence about agents, and Section 7.1 rates the strength of each conclusion accordingly.

6.2 The Promise: How Agents Can Improve Decisions

The strongest case for agentic commerce begins with Simon's insight about limited attention. A human shopper comparing running shoes might check three or four websites before fatigue sets in. An agent can, in principle, compare dozens of sellers in seconds, including shipping costs, return policies, and delivery dates that people often overlook. Research on online decision aids found that tools which screened and sorted options reduced search effort and improved the match between shoppers' preferences and the products they chose (Häubl and Trifts 2000). If agents perform this screening honestly, they could make careful comparison available to people who would never have done it themselves.

Agents may also find lower prices. Price-tracking services have existed for years, but they still require a person to notice an alert and complete the purchase in time. "Human-not-present" payments, which AP2 and the card networks are designed to support, could let a consumer set a rule such as "buy this jacket if it drops below \$60" and have the agent act at the right moment (Google Cloud 2025). For goods with fluctuating prices, such as airline tickets and electronics, this could produce real savings.

A third and often overlooked benefit is that agents can enforce rules people set for themselves. Behavioral economists have long observed that people struggle with self-control and can benefit from arrangements that bind their future behavior (Thaler and Sunstein 2008). An agent instructed to keep weekly grocery spending under \$75, or never to buy from a particular category, could act as a commitment device that never gets tired or tempted. The research on defaults that explains why unwanted subscriptions persist

also suggests that a well-chosen default, such as an agent that asks before any purchase over \$50, could be equally sticky in the consumer's favor (Madrian and Shea 2001; Johnson and Goldstein 2003). In that sense, an agent could restore some of the friction that one-click buying removed, but in a targeted way. Finally, agents could help people for whom online shopping is genuinely difficult. Much of the web remains poorly accessible: WebAIM's annual analysis of the top one million home pages found detectable accessibility failures on more than 94 percent of them in 2025 (WebAIM 2025). People with visual or motor disabilities, older adults unfamiliar with complicated checkout pages, and non-native speakers navigating websites in an unfamiliar language could all benefit from an agent that handles the mechanics of a purchase while they focus on what they want. Amazon has presented Buy for Me as a time-saver that spares customers from navigating unfamiliar brand websites (Amazon 2025).

These benefits are plausible, but independent researchers have tested few of them. Most of the evidence that agents save time or money comes from company announcements and consultant forecasts. The documentary analysis of product safeguards in Section 7.3 offers one check on how much of this promise current products deliver. Independent tests of AI shopping tools, and survey data on how much time they actually save, are needed before the promise can be treated as fact.

6.3 The Risks: How Agents Can Weaken Decisions

The same features that make agents convenient also create risks. The first is reduced awareness. If the pain of paying acts as a brake, and if card payments already weaken that brake (Prelec and Loewenstein 1998; Prelec and Simester 2001), an agent that completes purchases in the background may weaken it further. A consumer who never sees a checkout page never experiences the small moment of reckoning when a total appears on screen. Soman's (2001) finding that card users recall their spending less accurately suggests that the more a purchase happens out of sight, the less people remember about it. This reasoning underlies the review's second proposition.

The second risk is automation bias. An agent typically presents its result as a finished decision: "I found these shoes and placed your order." Research on automated aids suggests that people tend to accept such outputs without checking alternatives, especially when they are busy or the system has been reliable in the past (Skitka, Mosier, and Burdick 1999; Parasuraman and Manzey 2010). This creates a paradox. The more competent agents become, the more likely users are to stop reviewing their choices, which means the errors that remain are less likely to be caught.

And errors do occur. In February 2025, Washington Post technology columnist Geoffrey Fowler tested OpenAI's Operator by asking it to find inexpensive eggs for delivery. Instead of checking with him first, the agent completed a purchase of a dozen eggs that, with fees and a tip, cost about \$31, even though OpenAI had said Operator would ask for confirmation before purchases (Fowler 2025). The amount was small, but the episode shows how a confirmation safeguard can fail in practice. Other likely errors are mundane: the wrong size, the wrong quantity, a misread unit price, or a product that technically matches the instructions but not what the person meant.

Security is a fourth and more serious risk. Because agents read web pages to decide what to do, they can be manipulated by text hidden in those pages, a technique known as indirect prompt injection (Greshake et al. 2023). Benchmark studies have found that tool-using AI agents are frequently vulnerable to such attacks (Zhan et al. 2024; Debenedetti et al. 2024). Real-world tests have confirmed the danger. In August 2025, researchers at the browser company Brave showed that hidden instructions in a Reddit comment could direct Perplexity's Comet browser agent to retrieve a user's email address and a one-time login code (Chaikin and Sahib 2025). The same month, the security firm Guardio Labs reported that, in its tests, an

AI browser agent completed a purchase on a fake shopping website, filling in saved payment details without flagging that the store was fraudulent (Guardio Labs 2025).

AI developers themselves acknowledge the problem. Anthropic reported that, in internal testing of an experimental Claude browser extension, prompt injection attacks succeeded 23.6 percent of the time before new defenses were added and 11.2 percent afterward: a substantial improvement, but far from zero (Anthropic 2025). When OpenAI launched its ChatGPT Atlas browser in October 2025, the company's chief information security officer described prompt injection as "a frontier, unsolved security problem" (Stuckey 2025). An agent that holds a payment credential and can be redirected by a stranger's hidden text is an attractive target for scammers.

Privacy is a fifth risk. To shop on someone's behalf, an agent needs payment credentials, shipping addresses, and often a history of past purchases that reveals a great deal about a person's health, finances, and habits. Tokenization reduces the danger of card numbers being stolen, because the agent never sees the real number (Visa 2025), but it does nothing to limit how much behavioral data the agent's operator collects. The more useful an agent becomes by learning a person's preferences, the more sensitive the information it holds.

These risks reinforce one another. A consumer who trusts an agent, does not review its choices, and does not closely track spending is precisely the consumer least likely to notice a fraudulent or mistaken purchase. Each risk might be manageable on its own; combined, they describe a pattern in which money moves with less human attention at every stage. News reports of agent mistakes and security research on prompt injection, both of which are growing quickly, are the main evidence currently available, and both suggest that today's safeguards are still works in progress.

6.4 Who Does the Agent Work For?

Economists describe relationships in which one party acts on behalf of another as principal-agent relationships, and they have long recognized that agents may pursue their own interests at the principal's expense (Jensen and Meckling 1976). An AI shopping agent has at least three possible masters: the consumer who gives it instructions, the merchants whose products it selects, and the company that built and operates it. Researchers at Google DeepMind have argued that advanced AI assistants raise exactly this concern, because an assistant's helpfulness to its user can conflict with the commercial interests of its developer or of third parties (Gabriel et al. 2024).

Revenue models make the concern concrete. OpenAI says that merchants pay a fee on completed Instant Checkout purchases, while the service is free for users and does not change the prices they pay (OpenAI 2025c). The company states that product results are organic and unsponsored and are ranked by relevance. Yet the same announcement lists, among the factors ChatGPT considers when choosing between merchants that sell the same item, whether Instant Checkout is enabled (OpenAI 2025c). A merchant that has joined the fee-paying program may therefore be favored when other factors are equal. This is not necessarily improper, since a smooth checkout has real value to users, but it shows how an agent's choices can be shaped by the platform's business relationships in ways a user would not see. OpenAI's January 2026 announcement that it would begin testing labeled advertisements in ChatGPT, kept separate from its answers, sharpens the question of how commercial and conversational content will coexist (OpenAI 2026). Precedents from search engines are instructive. In 2017, the European Commission fined Google €2.42 billion for favoring its own comparison-shopping service in search results (European Commission 2017), a decision the EU's Court of Justice upheld in 2024 (*Google and Alphabet v. Commission*, Case C-48/22 P). In 2013, FTC staff wrote to search engine companies reminding them that paid results must be clearly

distinguishable from natural results (Federal Trade Commission 2013). The principle that advertising must be identifiable has not yet been clearly applied to the single answer an AI agent returns.

That single answer is itself a change in choice architecture. Thaler and Sunstein (2008) showed that the way options are presented shapes what people pick. An agent that returns one product instead of a list is the most extreme form of choice architecture, because the person sees only one option. Early experimental research suggests that AI agents are themselves sensitive to presentation. Researchers at Columbia University built a simulated shopping site and found that the products AI agents chose shifted with where items appeared on the page and with labels such as “sponsored” or platform endorsements, and that different AI models responded differently (Allouah et al. 2025). Microsoft researchers, studying a simulated marketplace of buying and selling agents, found that buyer agents showed a strong bias toward the first acceptable offer and could be manipulated by sellers (Bansal et al. 2025). If agents can be steered by position and labels, then whoever controls the page can influence the purchase.

The conflict between Amazon and Perplexity shows the stakes from the merchant side. In November 2025, Amazon sued Perplexity in federal court in California, alleging that Perplexity’s Comet browser agent shopped on Amazon’s site while disguising itself as a human user (*Amazon.com Services LLC v. Perplexity AI, Inc.*, N.D. Cal., filed November 4, 2025). Perplexity responded that users have the right to choose their own AI assistants. The dispute is partly about who controls the customer relationship, and with it, which products customers see.

Personalized pricing adds a further layer. Measurement studies have documented that some e-commerce sites show different prices or different product rankings to different users (Hannak et al. 2014). In 2024, the FTC ordered eight companies to provide information about “surveillance pricing,” the use of personal data to set individualized prices (Federal Trade Commission 2024), and in 2025 New York enacted a law requiring businesses to disclose when a price was set by an algorithm using a consumer’s personal data. An agent working for the consumer could help detect such pricing by comparing offers across sellers; an agent working for a seller could help impose it. Company disclosure policies on sponsored results and fees are therefore essential evidence, and the policies available so far vary in how clearly they answer the question of whom the agent serves.

6.5 Rules and Responsibility

When an agent buys the wrong item, someone must bear the cost: the consumer, the AI company, the merchant, or the bank. Existing law offers partial answers, and its gaps are revealing.

Payment law is the natural starting point. The federal protections described in Section 4.6 apply to “unauthorized” transactions, and the definitions of that word were not written with AI in mind. Regulation Z defines unauthorized credit card use as use by a person “who does not have actual, implied, or apparent authority” and from which the cardholder receives no benefit (12 C.F.R. § 1026.12(b)(1)(i)). Regulation E goes further: it excludes from its definition of an unauthorized transfer any transfer made by a person to whom the consumer furnished the means of access, unless the consumer has told the bank that this person is no longer authorized (12 C.F.R. § 1005.2(m)). A consumer who connects a card to an AI agent has arguably given it authority. Under that reading, an agent’s mistaken purchase may not be an unauthorized transaction at all, and the familiar \$50 liability caps may not apply.

A consumer might have better luck treating the problem as a billing error. The Fair Credit Billing Act lets credit card holders dispute charges for goods that were not accepted or not delivered in accordance with the agreement made at the time of the transaction (15 U.S.C. § 1666(b)). Whether an agent’s misreading

of an instruction fits that category has not been tested, and debit card users, who include many young people, do not have the same billing-error rights for purchases.

Contract law offers little more. The Uniform Electronic Transactions Act (UETA), adopted in 49 states, recognizes that contracts can be formed by “electronic agents” acting without human review (UETA § 14). Its error rule allows an individual to undo a transaction caused by a mistake made while dealing with another party’s electronic agent, but only if that agent gave no opportunity to prevent or correct the error (UETA § 10). That rule was designed for a person who mistypes on a merchant’s automated website, not for a consumer whose own agent makes the mistake.

Agency law, which governs human representatives, does not fill the gap either. The Restatement (Third) of Agency, an influential summary of US agency law, states that a computer program is not capable of acting as a principal or an agent in the legal sense (Restatement (Third) of Agency § 1.04 cmt. e (2006)). An AI “agent” therefore owes no legal duty of loyalty to the person it serves. Its behavior is governed instead by the terms of service the user accepted, which typically limit the developer’s liability. Legal scholars have begun to argue that principal–agent ideas still offer a useful framework for governing AI agents, even if the law does not yet treat them as agents (Kolt 2025).

Outside the United States, the European Union’s revised Product Liability Directive, adopted in 2024, explicitly treats software, including AI systems, as a product, so that defective AI can trigger liability without proof of fault (Directive (EU) 2024/2853). The directive, however, covers personal injury, property damage, and data loss, not purely financial losses such as an unwanted purchase.

The main US regulator for payments is in flux. The CFPB warned in 2023 that poorly designed chatbots in banking could fail consumers and break federal law (Consumer Financial Protection Bureau 2023). Since 2025, however, the bureau’s capacity has been sharply reduced, and Congress used the Congressional Review Act to overturn a 2024 CFPB rule that would have extended its supervision to large digital payment apps. In the near term, the agency most directly responsible for payment protections seems unlikely to lead on agentic commerce, leaving the issue to the FTC, the states, card network rules, and private contracts.

Industry is filling some gaps itself. Network tokens can be restricted by merchant and amount, and AP2’s signed mandates are designed to create a record of exactly what the user approved, which could help a bank decide whether a disputed purchase matched the user’s instructions (Google Cloud 2025; Mastercard 2025). These consent tools (spending limits, approval before payment, and easy refunds) are the most practical protections available today. But private standards can change at a company’s discretion and differ across products. Clear public rules about who pays for agent errors would give consumers more dependable protection than technical protocols they cannot see.

6.6 Young People and Financial Literacy

Teenagers and young adults may be affected differently from older consumers for two reasons: they are often early adopters of new AI tools, and they are still forming the financial habits that agents might replace. Financial literacy in the United States is uneven. Annamaria Lusardi and Olivia Mitchell (2014) document that large shares of adults cannot correctly answer basic questions about interest, inflation, and risk, and young people score especially low (Lusardi, Mitchell, and Curto 2010). Internationally, the OECD’s PISA 2022 assessment found that, on average across participating OECD countries, about 18 percent of 15-year-old students did not reach the baseline level of financial literacy (OECD 2024).

Education can make a difference. A study of states that introduced personal finance graduation requirements found improvements in young adults’ credit outcomes after the requirements took effect

(Urban et al. 2020). A growing number of states have adopted such requirements; Next Gen Personal Finance, which tracks state policies, reported that more than half of US states had guaranteed a personal finance course for high school students by 2025 (Next Gen Personal Finance 2025).

Agentic commerce raises a new question for this education: if AI handles comparison shopping and budgeting, do young people still learn those skills? Abilities such as reading unit prices or keeping track of spending develop through practice. An agent that compares options invisibly may produce a good purchase while removing the learning experience. The concern resembles long-running debates about calculators in math class: a tool is valuable once a skill is understood, but it may prevent the skill from forming if introduced too early.

Precedents from children’s in-app purchases show how easily automated payment systems can be misused by, or against, young users. In 2014, Apple agreed to refund at least \$32.5 million to settle FTC charges that it billed parents for in-app purchases their children made without consent (Federal Trade Commission 2014). In 2022, Epic Games agreed to pay \$245 million over design tactics that the FTC said led players of Fortnite, many of them young, into unwanted purchases (Federal Trade Commission 2022). Agentic purchasing could create similar risks if parental controls, age limits, and spending caps are not built in from the start. The EU AI Act’s ban on systems that exploit age-related vulnerabilities (Regulation (EU) 2024/1689, art. 5) points to the kind of protection young users may need.

For these reasons, schools may need to teach what could be called “AI financial literacy.” This would go beyond budgeting and price comparison to include setting an agent’s spending limits, checking its choices, recognizing when it may be steered by sponsorships, and disputing a mistaken charge. Such instruction treats the agent as a tool to be supervised rather than a replacement for judgment, which is the stance the research on appropriate reliance recommends (Lee and See 2004).

7. Discussion

The review finds strong evidence that less visible payment weakens spending restraint and that people over-rely on automation, but only limited and mostly indirect evidence about how AI shopping agents affect real consumers. This section answers each review question, evaluates the two propositions, examines the safeguards that current products document, and identifies gaps in the evidence.

7.1 Answers to the Review Questions

Table 4 summarizes the answer to each sub-question and rates the strength of the evidence behind it. *Strong* means consistent findings from several peer-reviewed studies or from the text of the law itself. *Moderate* means consistent findings from fewer or less rigorous studies, or from strong research that applies to agents only by analogy. *Limited* means early, anecdotal, or mainly grey-literature evidence.

Table 4. Summary of evidence by review question

Review question	What the evidence shows	Strength of evidence
1. Which purchases will people delegate?	Reliance on automation rises when errors seem cheap and falls after visible mistakes; no included study compares willingness to delegate AI purchases by price or category	Moderate, by analogy

2. Do agent users spend more or track spending less?	Less visible payment raises willingness to pay and lowers recall of spending; agents can remove the checkout step entirely	Strong for payment friction; limited for agents directly
3. Whose interests does the agent serve?	Fee arrangements and ranking factors can favor partner merchants; simulated agents are swayed by page position, labels, and first offers	Moderate
4. Who is responsible for mistakes and scams?	US payment law protects against unauthorized use, but an authorized agent's error may fall outside it; agents remain vulnerable to prompt injection	Strong for the legal gap; moderate for security risks
5. Are young people affected differently?	Young people show low financial literacy and benefit from education; automated payment systems have repeatedly been misused by or against children	Moderate; no included studies of agents and teenagers

7.2 Evaluating the Propositions

Proposition 1 receives partial support. Research on trust in automation consistently finds that reliance depends on perceived risk and on whether users can correct the system (Lee and See 2004; Dietvorst, Simmons, and Massey 2018), which fits the expectation that people prefer to delegate low-stakes, routine purchases. Current product designs also fit it: all three tools examined in Section 7.3 still ask for approval of each purchase, which suggests that companies expect consumers to want control over anything consequential. However, none of the included studies directly measures willingness to delegate purchases of different prices or types, so the proposition remains an inference rather than a finding.

Proposition 2 receives stronger, though still indirect, support. Studies spanning three decades show that the less visible a payment is, the more people are willing to spend and the less accurately they remember spending (Feinberg 1986; Prelec and Simester 2001; Soman 2001). Because an agent can remove even the moment of seeing a total, the mechanism these studies identify should operate at least as strongly with agents. Whether it actually does is an empirical question the existing literature has not yet answered.

7.3 Safeguards in Current Products: A Documentary Analysis

To see how the safeguards discussed in the literature appear in practice, the review extracted information from company documentation for three tools that represent the three models of agentic commerce identified in Section 4.5. Table 5 summarizes what each company's launch materials say about its safeguards.

Table 5. Safeguards in three agentic shopping tools, as described in company launch materials (2025)

Feature	ChatGPT agent	ChatGPT Instant Checkout	Amazon Buy for Me
Model	General-purpose browsing agent	In-chat checkout through the Agentic Commerce Protocol	Retailer-run agent inside the Amazon Shopping app

Who selects the product	The agent can search and propose items on any website, based on the user's instructions	ChatGPT surfaces products; the user picks one in the conversation	The user picks an item from another brand's site shown in Amazon search
Confirmation before payment	Asks permission before consequential actions such as purchases; the user can take over the browser	The user confirms the order, shipping, and payment details in the chat	The user taps a "Buy for Me" button to confirm
Payment handling	The user enters or approves payment details, which the agent is designed not to handle alone	A payment token is passed to the merchant, which processes the payment	Amazon securely passes encrypted billing details to the brand's site
User-set spending limit	None described	None described	None described
Returns and refunds	Handled by whichever merchant sold the item	Handled by the merchant, which remains the seller of record	Handled by the brand, not Amazon
Paid placement disclosure	Depends on the websites the agent visits	Merchants pay a fee on sales; results described as unsponsored, but Instant Checkout availability is a ranking factor	Not addressed in the launch announcement

Sources: OpenAI 2025b; OpenAI 2025c; Amazon 2025.

Three patterns stand out. First, current consumer products are more cautious than the introduction's scenario. Each requires the user to approve a specific purchase, so the fully autonomous shopper described in Section 1.1 is not yet the default. The infrastructure for that next step, however, already exists: AP2's "human-not-present" mandates and the card networks' agent tokens are designed precisely for purchases made without real-time approval (Google Cloud 2025; Mastercard 2025; Visa 2025). Second, none of the three tools describes a budget cap that users can set, even though spending limits are among the most commonly proposed safeguards. Third, responsibility is split. In every case, the company whose AI guided the purchase is different from the company the consumer must contact for a refund. This split supports the concern raised in Section 6.5 that accountability for agent errors is unclear.

These findings come with an important caveat. They are based on company descriptions rather than hands-on testing with real money, and the products change frequently. Fowler's (2025) experience with Operator shows that a stated confirmation policy does not guarantee that confirmation always happens.

7.4 Gaps in the Evidence

The review reveals four gaps. First, none of the included sources tracks real consumers' spending before and after they begin using AI agents, so claims that agents save money or encourage overspending both

remain untested. Second, evidence that agents favor sponsored or fee-paying merchants comes from simulations and company policy statements rather than independent audits of real products. Third, the review found no court decision or regulatory guidance addressing whether an agent's mistaken purchase counts as unauthorized under Regulation E or Regulation Z. Fourth, no included study examines how teenagers use AI shopping tools or whether doing so affects their financial learning. Section 9.4 proposes research to fill these gaps.

8. Ethical Considerations, Limitations, and Counterarguments

Naming a study's weaknesses openly makes its argument more convincing, not less. This section describes the ethical standards the review follows, where its evidence falls short, and how the thesis holds up against the strongest objections.

8.1 Ethical Considerations

Because an extended literature review uses only published sources, it involves no human participants, so the consent and data-protection procedures required for surveys and interviews do not apply. A review carries its own ethical obligations, however. Sources must be represented accurately, including findings that cut against the thesis, such as the evidence for algorithm aversion and the weak average effect of choice overload. Every citation must be checked against the original source rather than copied from another paper's reference list. Conflicts of interest in the evidence must be reported, which is why company announcements are labeled as promotional throughout. Finally, Section 5 describes how sources were found and selected so that readers can judge how the conclusions were reached.

8.2 Limitations

The review has several limitations. An extended review is less exhaustive than a systematic review: because sources were selected by judgment rather than screened against a fixed protocol, relevant studies may have been missed, and the selection may reflect the author's perspective despite the deliberate search for contrary evidence. The review includes only English-language sources available through open access or a school library, which may omit research published elsewhere or behind paywalls. Published research tends to over-represent striking results, a problem known as publication bias, which may inflate the apparent size of effects such as the credit card premium. Much of the direct evidence about AI agents comes from preprints, simulations, and company or news sources rather than peer-reviewed studies of real consumers. Finally, agentic commerce is changing so quickly that products and policies described here may have changed by the time this paper is read.

8.3 Counterarguments

The first objection is that agentic commerce is "just one-click buying again." On this view, consumers adapted to saved cards and one-click ordering, and they will adapt to agents. The response is that the two differ in kind, not just degree. One-click buying removed friction from paying, but the consumer still chose the product. An agent removes friction from choosing, so the consumer may never see the alternatives at all. Agents also introduce risks that one-click buying never carried, such as prompt injection, in which a stranger's hidden text can redirect a purchase (Greshake et al. 2023).

The second objection is that people already overspend, and an AI with a budget cap could help them. This paper agrees in part. Section 6.2 showed that agents could serve as commitment devices, enforcing limits that people struggle to keep on their own. But the outcome depends on design and defaults. None of the three tools examined in Section 7.3 describes a user-set spending cap, and research on defaults shows that

safeguards people must actively switch on are used far less than safeguards that are on by default (Madrian and Shea 2001; Johnson and Goldstein 2003). A budget cap helps only if it exists and is on.

The third objection is that consumers are free to turn agents off. Freedom to opt out, however, does not mean people will use it. Johnson and Goldstein (2003) found that organ donor registration rates differed dramatically between countries largely because of whether donation was the default, not because people's values differed. Convenience exerts a similar pull. As agentic checkout becomes embedded in the platforms people already use, declining it may mean giving up speed, discounts, or features. The real question is not whether consumers can turn agents off, but whether the default settings protect those who never do.

9. Conclusion

Agentic AI can make shopping faster and sometimes cheaper, but it removes the pauses in which people compare and reconsider, and whether it helps or harms consumers depends on the safeguards built around it.

9.1 The Thesis Revisited

The evidence reviewed supports that thesis, with one important qualification. Decades of research on the pain of paying and on automation bias give strong grounds for expecting that purchases made out of sight will be noticed and questioned less. Direct evidence about AI agents, though limited, points the same way: agents make errors, can be manipulated by hidden instructions, and may be steered by commercial arrangements. The documentary analysis found that today's leading tools still require a human to approve each purchase, but that none describes a user-set spending limit and that responsibility for errors is split between the AI company and the merchant. The qualification is that the strongest evidence concerns human behavior in general rather than agents specifically, so part of the thesis rests on well-grounded inference rather than direct measurement.

9.2 Summary of Key Findings

On the benefits, agents can compare more options than any person, act on price drops, enforce budgets, and help people who find online shopping difficult, though independent evidence for these benefits remains thin. On the risks, agents weaken the pain of paying, invite automation bias, make errors, and can be hijacked through prompt injection, as security researchers demonstrated repeatedly in 2025. On conflicts of interest, an agent can serve its user, its merchants, or its developer, and fees, advertising, and ranking rules can tilt its single recommendation in ways users cannot see. On responsibility, US payment protections were written for stolen cards, not for authorized agents that make mistakes, leaving a gap that private protocols only partly fill. On young people, agents may handle the very tasks through which teenagers learn to manage money, which argues for teaching AI financial literacy alongside traditional budgeting.

9.3 Recommendations

The findings point to practical steps for four groups:

- **Consumers** should set spending limits wherever a tool allows them, keep purchase confirmation turned on, and check receipts and statements weekly.
- **Companies** should label sponsored or fee-influenced picks clearly, offer user-set budget caps that are on by default, and require explicit approval above a set dollar amount.
- **Policymakers** should clarify whether an agent's mistaken purchase counts as an unauthorized transaction or a billing error, and who pays when an agent is deceived by a scam.

- **Schools** should add AI tools to financial literacy lessons, teaching students to supervise agents rather than simply trust them.

9.4 Future Research

The gaps identified in Section 7.4 set the agenda for future research. A survey of students, parents, and teachers could measure directly which purchases people are willing to delegate and whether frequent delegators misjudge their spending, testing the two propositions that this review could evaluate only indirectly. A simple experiment could compare people who see only an AI's single recommendation with people who see the full list of options, measuring the price they pay and their confidence in the choice. Longer-term studies could track real spending as people adopt agents, and independent audits could test whether agents' recommendations favor sponsored or fee-paying merchants. Research focused on teenagers would show whether AI shopping tools help or hinder the development of financial skills.

9.5 Closing Thought

Return, finally, to the student who asked an AI for running shoes under \$80. Did the student really make the decision? In one sense, yes: the student decided to delegate, set the budget, and named the deadline. But the student did not choose the shoes, did not see what was given up, and may not remember the purchase a month later. A decision made through an agent remains truly one's own only when the person sets its boundaries in advance, can inspect the choice, and keeps the power to undo it. Without those safeguards, the student did not choose a pair of shoes. The student chose to let something else choose, and that something else may have had interests of its own.

References

1. *Statutes, regulations, court cases, patents, and EU legislation are cited in the text, following Chicago author-date practice.*
2. Adams, Richard J., Palie Smart, and Anne Sigismund Huff. 2017. "Shades of Grey: Guidelines for Working with the Grey Literature in Systematic Reviews for Management and Organizational Studies." *International Journal of Management Reviews* 19 (4): 432–54.
3. Adobe. 2025. "Adobe Analytics: Traffic to U.S. Retail Websites from Generative AI Sources Jumps 1,200 Percent." Adobe Newsroom, March 2025.
4. Allouah, Amine, Omar Besbes, Josué D. Figueroa, Yash Kanoria, and Akshit Kumar. 2025. "What Is Your AI Agent Buying? Evaluation, Implications, and Emerging Questions for Agentic E-Commerce." Preprint, arXiv, August 2025. arXiv:2508.02630.
5. Amazon. 2025. "Amazon Tests 'Buy for Me,' a New AI Shopping Agent That Lets Customers Buy from Other Brands' Websites." About Amazon, April 3, 2025.
6. Anthropic. 2025. "Piloting Claude for Chrome." Anthropic News, August 26, 2025. <https://www.anthropic.com/news/claude-for-chrome>.
7. Bansal, Gagan, et al. 2025. "Magentic Marketplace: An Open-Source Environment for Studying Agentic Markets." Preprint, Microsoft Research, November 2025.
8. Baumeister, Roy F., and Mark R. Leary. 1997. "Writing Narrative Literature Reviews." *Review of General Psychology* 1 (3): 311–20.
9. Chaikin, Artem, and Shivan Kaul Sahib. 2025. "Agentic Browser Security: Indirect Prompt Injection in Perplexity Comet." Brave Blog, August 20, 2025.
10. Consumer Financial Protection Bureau. 2023. *Chatbots in Consumer Finance*. Issue Spotlight. Washington, DC: Consumer Financial Protection Bureau, June 2023.

11. D'Acunto, Francesco, Nagpurnanand Prabhala, and Alberto G. Rossi. 2019. "The Promises and Pitfalls of Robo-Advising." *Review of Financial Studies* 32 (5): 1983–2020.
12. Debenedetti, Edoardo, Jie Zhang, Mislav Balunović, Luca Beurer-Kellner, Marc Fischer, and Florian Tramèr. 2024. "AgentDojo: A Dynamic Environment to Evaluate Prompt Injection Attacks and Defenses for LLM Agents." In *Advances in Neural Information Processing Systems 37: Datasets and Benchmarks Track*.
13. Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. 2015. "Algorithm Aversion: People Erroneously Avoid Algorithms after Seeing Them Err." *Journal of Experimental Psychology: General* 144 (1): 114–26.
14. Dietvorst, Berkeley J., Joseph P. Simmons, and Cade Massey. 2018. "Overcoming Algorithm Aversion: People Will Use Imperfect Algorithms If They Can (Even Slightly) Modify Them." *Management Science* 64 (3): 1155–70.
15. Engel, James F., David T. Kollat, and Roger D. Blackwell. 1968. *Consumer Behavior*. New York: Holt, Rinehart and Winston.
16. European Commission. 2017. "Antitrust: Commission Fines Google €2.42 Billion for Abusing Dominance as Search Engine by Giving Illegal Advantage to Own Comparison Shopping Service." Press release, June 27, 2017.
17. Federal Trade Commission. 2013. "FTC Consumer Protection Staff Updates Agency's Guidance to Search Engine Industry on the Need to Distinguish Between Advertisements and Search Results." Press release, June 25, 2013.
18. Federal Trade Commission. 2014. "Apple Inc. Will Provide Full Consumer Refunds of at Least \$32.5 Million to Settle FTC Complaint It Charged for Kids' In-App Purchases Without Parental Consent." Press release, January 15, 2014.
19. Federal Trade Commission. 2022. "Fortnite Video Game Maker Epic Games to Pay More Than Half a Billion Dollars over FTC Allegations of Privacy Violations and Unwanted Charges." Press release, December 19, 2022.
20. Federal Trade Commission. 2024. "FTC Issues Orders to Eight Companies Seeking Information on Surveillance Pricing." Press release, July 23, 2024.
21. Federal Trade Commission. 2025. "FTC Secures Historic \$2.5 Billion Settlement Against Amazon." Press release, September 25, 2025.
22. Feinberg, Richard A. 1986. "Credit Cards as Spending Facilitating Stimuli: A Conditioning Interpretation." *Journal of Consumer Research* 13 (3): 348–56.
23. Fowler, Geoffrey A. 2025. "I Let ChatGPT's New Agent Manage My Life. It Spent \$31 on a Dozen Eggs." *Washington Post*, February 7, 2025.
24. Gabriel, Iason, Arianna Manzini, Geoff Keeling, et al. 2024. "The Ethics of Advanced AI Assistants." Preprint, arXiv, April 2024. arXiv:2404.16244.
25. Garousi, Vahid, Michael Felderer, and Mika V. Mäntylä. 2019. "Guidelines for Including Grey Literature and Conducting Multivocal Literature Reviews in Software Engineering." *Information and Software Technology* 106: 101–21.
26. Google Cloud. 2025. "Announcing Agent Payments Protocol (AP2)." Google Cloud Blog, September 16, 2025.
27. Grant, Maria J., and Andrew Booth. 2009. "A Typology of Reviews: An Analysis of 14 Review Types and Associated Methodologies." *Health Information and Libraries Journal* 26 (2): 91–108.

28. Greshake, Kai, Sahar Abdelnabi, Shailesh Mishra, Christoph Endres, Thorsten Holz, and Mario Fritz. 2023. "Not What You've Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection." In *Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security*, 79–90. New York: Association for Computing Machinery.
29. Guardio Labs. 2025. "Scamlexity: We Put Agentic AI Browsers to the Test—They Clicked, They Paid, They Failed." Guardio Labs Blog, August 2025.
30. Hannak, Aniko, Gary Soeller, David Lazer, Alan Mislove, and Christo Wilson. 2014. "Measuring Price Discrimination and Steering on E-commerce Web Sites." In *Proceedings of the 2014 Internet Measurement Conference*, 305–18. New York: Association for Computing Machinery.
31. Häubl, Gerald, and Valerie Trifts. 2000. "Consumer Decision Making in Online Shopping Environments: The Effects of Interactive Decision Aids." *Marketing Science* 19 (1): 4–21.
32. Iyengar, Sheena S., and Mark R. Lepper. 2000. "When Choice Is Demotivating: Can One Desire Too Much of a Good Thing?" *Journal of Personality and Social Psychology* 79 (6): 995–1006.
33. Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics* 3 (4): 305–60.
34. Johnson, Eric J., and Daniel Goldstein. 2003. "Do Defaults Save Lives?" *Science* 302 (5649): 1338–39.
35. Kahneman, Daniel. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
36. Knutson, Brian, Scott Rick, G. Elliott Wimmer, Drazen Prelec, and George Loewenstein. 2007. "Neural Predictors of Purchases." *Neuron* 53 (1): 147–56.
37. Kolt, Noam. 2025. "Governing AI Agents." *Notre Dame Law Review* 101.
38. Kotler, Philip, and Kevin Lane Keller. 2016. *Marketing Management*. 15th ed. Boston: Pearson.
39. Lee, John D., and Katrina A. See. 2004. "Trust in Automation: Designing for Appropriate Reliance." *Human Factors* 46 (1): 50–80.
40. Logg, Jennifer M., Julia A. Minson, and Don A. Moore. 2019. "Algorithm Appreciation: People Prefer Algorithmic to Human Judgment." *Organizational Behavior and Human Decision Processes* 151: 90–103.
41. Lusardi, Annamaria, and Olivia S. Mitchell. 2014. "The Economic Importance of Financial Literacy: Theory and Evidence." *Journal of Economic Literature* 52 (1): 5–44.
42. Lusardi, Annamaria, Olivia S. Mitchell, and Vilsa Curto. 2010. "Financial Literacy among the Young." *Journal of Consumer Affairs* 44 (2): 358–80.
43. Madrian, Brigitte C., and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116 (4): 1149–87.
44. Mastercard. 2025. "Mastercard Unveils Agent Pay, Pioneering Agentic Payments Technology to Power Commerce in the Age of AI." Press release, April 29, 2025.
45. McKinsey & Company. 2025. *The Agentic Commerce Opportunity: How AI Agents Are Ushering in a New Era for Consumers and Merchants*. McKinsey & Company, October 2025.
46. Next Gen Personal Finance. 2025. "Personal Finance Guarantee Tracker." Next Gen Personal Finance.
47. OECD. 2024. *PISA 2022 Results (Volume IV): How Financially Smart Are Students?* Paris: OECD Publishing.
48. OpenAI. 2025a. "Introducing Operator." January 23, 2025. <https://openai.com/index/introducing-operator/>.

49. OpenAI. 2025b. "Introducing ChatGPT Agent: Bridging Research and Action." July 17, 2025. <https://openai.com/index/introducing-chatgpt-agent/>.
50. OpenAI. 2025c. "Buy It in ChatGPT: Instant Checkout and the Agentic Commerce Protocol." September 29, 2025. <https://openai.com/index/buy-it-in-chatgpt/>.
51. OpenAI. 2026. "Our Approach to Advertising and Expanding Access to ChatGPT." January 2026.
52. Parasuraman, Raja, and Dietrich H. Manzey. 2010. "Complacency and Bias in Human Use of Automation: An Attentional Integration." *Human Factors* 52 (3): 381–410.
53. Parasuraman, Raja, and Victor Riley. 1997. "Humans and Automation: Use, Misuse, Disuse, Abuse." *Human Factors* 39 (2): 230–53.
54. Pew Research Center. 2025. "How Americans View AI and Its Impact on People and Society." September 17, 2025.
55. Prelec, Drazen, and George Loewenstein. 1998. "The Red and the Black: Mental Accounting of Savings and Debt." *Marketing Science* 17 (1): 4–28.
56. Prelec, Drazen, and Duncan Simester. 2001. "Always Leave Home Without It: A Further Investigation of the Credit-Card Effect on Willingness to Pay." *Marketing Letters* 12 (1): 5–12.
57. Scheibehenne, Benjamin, Rainer Greifeneder, and Peter M. Todd. 2010. "Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload." *Journal of Consumer Research* 37 (3): 409–25.
58. Shah, Avni M., Noah Eisenkraft, James R. Bettman, and Tanya L. Chartrand. 2016. "Paper or Plastic?: How We Pay Influences Post-Transaction Connection." *Journal of Consumer Research* 42 (5): 688–708.
59. Shavit, Yonadav, Sandhini Agarwal, Miles Brundage, et al. 2023. "Practices for Governing Agentic AI Systems." OpenAI white paper, December 2023.
60. Simon, Herbert A. 1955. "A Behavioral Model of Rational Choice." *Quarterly Journal of Economics* 69 (1): 99–118.
61. Skitka, Linda J., Kathleen L. Mosier, and Mark Burdick. 1999. "Does Automation Bias Decision-Making?" *International Journal of Human-Computer Studies* 51 (5): 991–1006.
62. Smith, Brent, and Greg Linden. 2017. "Two Decades of Recommender Systems at Amazon.com." *IEEE Internet Computing* 21 (3): 12–18.
63. Snyder, Hannah. 2019. "Literature Review as a Research Methodology: An Overview and Guidelines." *Journal of Business Research* 104: 333–39.
64. Soman, Dilip. 2001. "Effects of Payment Mechanism on Spending Behavior: The Role of Rehearsal and Immediacy of Payments." *Journal of Consumer Research* 27 (4): 460–74.
65. Stuckey, Dane. 2025. Post on X about prompt injection and ChatGPT Atlas. October 22, 2025.
66. Thaler, Richard H., and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
67. Torraco, Richard J. 2005. "Writing Integrative Literature Reviews: Guidelines and Examples." *Human Resource Development Review* 4 (3): 356–67.
68. Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185 (4157): 1124–31.
69. Tyndall, Jess. 2010. *AACODS Checklist*. Adelaide: Flinders University.

70. Urban, Carly, Maximilian Schmeiser, J. Michael Collins, and Alexandra Brown. 2020. "The Effects of High School Personal Financial Education Policies on Financial Behavior." *Economics of Education Review* 78: 101786.
71. Visa. 2025. "Find and Buy with AI: Visa Unveils New Era of Commerce." Press release, April 30, 2025.
72. WebAIM. 2025. *The WebAIM Million: The 2025 Report on the Accessibility of the Top 1,000,000 Home Pages*. Logan, UT: WebAIM.
73. Wohlin, Claes. 2014. "Guidelines for Snowballing in Systematic Literature Studies and a Replication in Software Engineering." In *Proceedings of the 18th International Conference on Evaluation and Assessment in Software Engineering*, article 38. New York: Association for Computing Machinery.
74. Xiao, Bo, and Izak Benbasat. 2007. "E-Commerce Product Recommendation Agents: Use, Characteristics, and Impact." *MIS Quarterly* 31 (1): 137–209.
75. Zhan, Qiusi, Zhixiang Liang, Zifan Ying, and Daniel Kang. 2024. "InjecAgent: Benchmarking Indirect Prompt Injections in Tool-Integrated Large Language Model Agents." In *Findings of the Association for Computational Linguistics: ACL 2024*. Bangkok: Association for Computational Linguistics.