



Prifysgol Cymru
Y Drindod Dewi Sant
University of Wales
Trinity Saint David

The Impact of AI Tools on Customer Support in Online Shopping in the UK

BA (Hons) Business Management

LCBB6014 Independent Project (Dissertation)

Student Name: Istvan Tarjani

Student ID: 2317622

Module Leader: Dr Dudley Ofori

Supervisor: Dr Lakshani Nissanga Mendis

Submission Date: 21 August 2026

A combined secondary-research and cross-sectional survey study.

The dissertation contains **9,208** words, counted from **Chapters 1–5**, excluding tables, diagrams, the table of contents, lists of tables and figures, references, the glossary, appendices, the title page and declaration.

August 2026

Declaration

I declare that this dissertation is my own work and that all sources used have been acknowledged in accordance with the required Harvard referencing style. This work has not been submitted, in whole or in part, for another academic award. Where digital or AI tools have assisted with the research or preparation of this dissertation, their use is disclosed in the Academic-Integrity Statement. Responsibility for the submitted work remains with me.

Student name: Istvan Tarjani

Student ID: 2317622

Signature: Istvan Tarjani

Date: 17.08.2026

Abstract

This dissertation evaluates the main AI tools encountered in customer support for UK online retail and assesses how those tools relate to service performance, advantages and constraints, and the described online-shopping experience. It integrates a structured review of the retained 40 source evidence corpus with findings from a cross-sectional online survey that received supervisor approval on 25 April 2026 before data collection. The primary investigation involved 147 eligible respondents recruited through Discord, Facebook and X.com(twitter) between 3 and 31 July 2026. All respondents consented, were aged 18 or over and resident in the UK, and reported online-shopping exposure to AI support during the previous 12 months. The questionnaire measured tool and task, journey stage, response speed, reliability, convenience, effectiveness, service quality, satisfaction, trust, privacy concern, ease, independent completion, retailer expectations, human access and improvement priorities.

Counts, percentages, complete ordinal distributions and medians provide the main results. Wilson 95% confidence intervals quantify uncertainty around selected proportions, and percentile bootstrap intervals based on 5,000 resamples accompany the eleven pre-specified Spearman coefficients. Exploratory tool- and task-group comparisons use pooled categories to avoid reporting unstable single-digit cells. Text-based chatbots were identified by 80 respondents (54.4%). Checking an order or delivery and returning or exchanging a product were the joint-leading tasks, with 32 responses each. Positive ratings were 71.4% for speed, 58.5% for reliability, 68.7% for convenience, 55.8% for effectiveness, 59.2% for service quality and 55.1% for satisfaction. Strong trust was reported by 36.7%, high privacy concern by 36.1%, and access to a human adviser when AI could not resolve a problem was very or extremely important to 83.0%. The largest examined Spearman coefficient was .225, so the outcomes do not support a strong, representative or causal account of AI effects.

The combined evidence presents speed, convenience, ease and independent completion as potential advantages, while mixed resolution, limited strong trust, privacy concern, uneven participation and inconsistent human recovery remain important limitations. The described online-shopping experience depended on task relevance, service performance, expectations, successful resolution and access to accountable human assistance. The dissertation recommends task-bounded automation, separate performance and satisfaction monitoring, proportionate data use, inclusive testing and effective escalation. The conclusions apply to the achieved convenience sample and retained evidence corpus, not to all UK consumers or retailers.

Contents

Declaration	2
Abstract	3
List of Tables	6
List of Figures	7
Chapter 1: Introduction	8
1.1 Background to the Study	8
1.2 Problem Statement	8
1.3 Research Aim, Objectives and Questions	9
1.4 Significance of the Study	9
1.5 Scope and Definitions	9
1.6 Structure of the Dissertation	9
1.7 Chapter Summary	10
Chapter 2: Review of Literature	10
2.1 Introduction	10
2.2 AI Tools in Customer Support	10
2.3 Chatbots, Interaction Quality and Human Presence	10
2.4 AI Support, Satisfaction and Trust	11
2.5 Service Quality, Speed and Resolution	12
2.6 Personalization and Automation	12
2.7 Privacy, Ethics and Responsible AI	12
2.8 Digital Inclusion and Vulnerable Customers	13
2.9 UK Online Retail Context	13
2.10 Theoretical Frameworks	14
2.11 The Hybrid Model for UK Online Retail	15
2.12 Research Gap and Chapter Summary	16
Chapter 3: Research Methods and Methodology	17
3.1 Introduction	17
3.2 Research Philosophy and Approach	17
3.3 Combined Research Design	17
3.4 Population, Sampling and Achieved Sample	17
3.5 Questionnaire and Measures	18
3.6 Secondary Research Procedure	18
3.7 Data Preparation, Coding and Quality Control	19
3.8 Analysis and Triangulation Strategy	19
3.9 Reliability, Validity and Ethics	20
3.10 Methodological Limitations	20
3.11 Chapter Summary	20
Chapter 4: Findings, Analysis and Discussion	21
4.1 Data Quality and Sample Profile	21
4.2 Main AI Tools in UK Online Retail Customer Support	21
4.3 Service Performance and Customer Satisfaction	23
4.4 Benefits and Limitations: Trust, Privacy and Digital Exclusion	25
4.5 Effects on the Online Shopping Experience	27
4.6 Practical Recommendations for Responsible and Effective AI Use	28
4.7 Integrated Discussion	28
4.8 Chapter Summary	30
Chapter 5: Conclusions, Limitations and Recommendations	31

5.1 Introduction.....	31
5.2 Answers to the Research Questions.....	31
5.3 Achievement of the Aim and Contribution	31
5.4 Recommendations for UK Online Retailers	32
5.5 Limitations	32
5.6 Future Research	33
5.7 Final Conclusion.....	33
References.....	34
Glossary of Terms.....	38
Appendices	39
Appendix A: Administered Questionnaire	39
Appendix B: Objective and Question Mapping	41
Appendix C: Coding and Analysis Protocol	42
Appendix D: Data-Quality Audit	42
Appendix E: Secondary-Evidence Search and Appraisal.....	42
Appendix F: Ethics Statement.....	45
Appendix G: Research Timeline	45
Appendix H: AI-Use and Academic-Integrity Statement	46

List of Tables

Table 2.1 Theoretical lenses and analytical use	14
Table 3.1 Objective, question and evidence mapping	18
Table 3.2 Coding and analysis rules	19
Table 4.1 Data quality and sample profile	21
Table 4.2 Leading tools, tasks and journey stages	21
Table 4.3 Full support-performance distributions	23
Table 4.4 Exploratory tool-group performance and satisfaction	25
Table 4.5 Exploratory task-group performance and satisfaction	25
Table 4.6 Satisfaction, trust, privacy, ease and expectation distributions	26
Table 4.7 Descriptive inclusion comparisons	26
Table 4.8 Human access, influence and improvement priorities	27
Table 4.9 Pre-specified Spearman associations with uncertainty intervals	29
Table 4.10 Integrated evidence and managerial implications	30
Table 5.1 Evidence-linked recommendations for UK online retailers	32
Table B1 Question-to-objective mapping	41
Table C1 Coding protocol	42
Table D1 Data-quality audit	42
Table E1 Secondary-evidence selection and appraisal	43
Table G1 Research timeline	45

List of Figures

Figure 2.1 Combined evidence framework for AI customer support 14

Figure 4.1 Reported AI support tools 22

Figure 4.2 Main customer-service tasks 23

Figure 4.3 Positive support-performance ratings..... 24

Figure 4.4 Satisfaction, trust, privacy, ease and independent completion 26

Figure 4.5 Human-access outcomes 28

Chapter 1: Introduction

1.1 Background to the Study

Online shopping is embedded in UK consumer activity. Office for National Statistics series show that internet sales remain a material part of retail spending, although the share changes across months and sectors (ONS, 2026a; ONS, 2026b; ONS, 2026c). This scale creates customer-support needs before payment, during fulfilment and after delivery. For this dissertation, text chatbots, virtual assistants, automated tracking and returns, recommendation functions, AI help search and agentic systems form an operational taxonomy. Conversational-agent research and current organizational evidence inform the categories, but do not establish their market prevalence in UK retail (Mariani, Hashemi and Wirtz, 2023; McKinsey & Company, 2025). Their presence, however, does not establish their value to customers.

AI support can respond immediately to routine enquiries and may extend service availability beyond staffed hours (Uzoka, Cadet and Ojukwu, 2024). A rapid reply may be useful for product information or order status, but speed has limited value if the answer is wrong, the system lacks authority or the customer cannot recover from failure. Prior work on AI-chatbot service quality separates responsiveness from reliability, assurance and problem-solving competence (Chen et al., 2022). E-commerce evidence similarly indicates that interaction design and perceived capability can affect experience and decision-making, but does not make every task or context equivalent (Chau et al., 2025).

The customer journey therefore provides the practical unit of analysis. Recommendation tools influence choice, tracking systems report fulfilment, and returns or complaint systems operate where money, rights or redress may be involved. The consequences of error differ across these tasks. This dissertation keeps speed, reliability, convenience, effectiveness, service quality, satisfaction, trust, privacy concern, ease, independent completion and human access analytically separate rather than treating them as a single quality score (Mariani, Hashemi and Wirtz, 2023; Chen et al., 2022; Ozuem et al., 2024).

UK adoption also occurs across uneven access, confidence and skills (Ofcom, 2026; Lloyds Banking Group, 2025). AI support can simplify a routine interaction, but opaque data use, inaccessible interfaces or a blocked escalation route may transfer effort and risk to the customer (Hasal et al., 2021; Liao and Sundar, 2022; Ozuem et al., 2024). Responsible evaluation must consequently consider who can use the channel, what happens when it fails, and whether a person with appropriate authority remains reachable.

1.2 Problem Statement

The evidence base does not establish consistent customer benefit across contexts. Much published work examines adoption intentions, banking, airports or non-UK markets, while commercial reports emphasise capability, scale and efficiency (Auer, Schlögl and Glowka, 2024; McKinsey & Company, 2025). Cross-sectional surveys frequently associate perceived quality, trust and satisfaction, but cannot demonstrate that AI caused the observed judgement (Nguyen and Le, 2025; Beyari, 2025). Operational measures such as contact deflection or first-response time can also obscure customer satisfaction, repeat contact, unresolved problems, digital exclusion and inaccessible recovery.

The evidence gap is therefore broader than whether consumers like chatbots. UK online retailers need to know which AI tools customers encounter, how speed and service quality compare with satisfaction, which advantages and constraints arise around trust, privacy and digital exclusion, and how support contributes to the wider online-shopping experience. These questions require both empirical description and critical secondary context. This investigation addresses them by integrating a 147-person survey with the retained 40-source evidence corpus.

The achieved sample is an online convenience sample. It can show patterns within the 147 responses, including weak or unexpected associations, but it cannot estimate the prevalence of an opinion among all UK shoppers. The cross-sectional design also cannot establish causal AI effects. These boundaries are carried through the assessment, guidance and conclusion (Bell, Bryman and Harley, 2022).

1.3 Research Aim, Objectives and Questions

The research aim is: To critically evaluate the use and reported effects of AI customer support in UK online retail, including service performance, customer satisfaction, benefits, limitations and the online-shopping experience, and to provide responsible and effective guidance for retailers.

The research objectives are:

1. To identify the main AI tools used by UK online retailers for customer support.
2. To assess how AI tools influence service performance, including response speed, service quality, and customer satisfaction.
3. To explore the benefits and limitations of AI support, including trust, privacy, and digital exclusion.
4. To examine how AI customer support affects the online shopping experience.
5. To provide practical recommendations for UK online retailers on responsible and effective AI use.

The corresponding research questions are:

RQ1: Which main AI customer-support tools are identified in the reviewed evidence and most frequently reported in respondents' recent experiences with UK online retailers?

RQ2: How do respondents' ratings characterise the influence of AI support on response speed, reliability, convenience, effectiveness, service quality and customer satisfaction?

RQ3: What benefits and limitations of AI support are shown by the combined evidence, particularly in relation to trust, privacy and digital exclusion?

RQ4: How is AI customer support associated with the reported online-shopping experience across journey stages, expectations, problem resolution and access to human assistance?

RQ5: What practical actions should UK online retailers take to use AI responsibly and effectively?

1.4 Significance of the Study

This investigation contributes by linking the AI tools reported in customer-support encounters to distinct performance, satisfaction and experience outcomes. For retailers, separating advantages from limitations supports more accountable decisions than adoption or speed metrics alone. Academically, the investigation examines whether concepts often connected in other sectors appear strongly related in a UK online-shopping sample. Socially, it treats trust, privacy, digital exclusion and access to human judgement as central elements of responsible and effective AI-supported service (Chen et al., 2022; Liao and Sundar, 2022).

1.5 Scope and Definitions

The scope is customer-facing AI encountered before, during or after an online purchase by UK adults. Service performance covers speed, reliability, convenience, effectiveness and service quality; satisfaction is the respondent's overall evaluation. Advantages and constraints are examined through relevance, usability, trust, privacy and participation barriers. Online-shopping experience is interpreted through journey stage, satisfaction, expectations, resolution and human access. These are stated indicators, not verified retailer-system audits or causal effects.

1.6 Structure of the Dissertation

Chapter 2 reviews the retained evidence corpus and analytical framework. Chapter 3 explains the survey, secondary review, revised mapping, assessment, ethics and limitations. Chapter 4 reports findings by objective. Chapter 5 answers the research questions and presents responsible, effective retailer recommendations. The appendices preserve the questionnaire, mapping, coding and quality controls.

1.7 Chapter Summary

This chapter has framed AI support as a conditional part of online shopping. Value depends on task fit, dependable performance, acceptable trust and privacy, inclusive access and effective recovery. Chapter 2 examines those conditions (Chen et al., 2022; Ozuem et al., 2024).

Chapter 2: Review of Literature

2.1 Introduction

This chapter establishes the concepts used for the revised objectives. It compares prior work on AI tools, service performance, satisfaction, online-shopping experience, trust, privacy, digital exclusion and recovery across methods, sectors and countries. Four theoretical lenses are combined without assuming a single scale or causal effect (Chen et al., 2022; Liao and Sundar, 2022).

AI support is classified by task, data access and authority: recommendation assistants influence choice, tracking tools report fulfilment, and return or complaint systems operate after failure. The taxonomy draws on conversational-agent reviews (Mariani, Hashemi and Wirtz, 2023; Nicolescu and Tudorache, 2022), but does not establish retailer prevalence, equal capability or uniform effects.

2.2 AI Tools in Customer Support

Customer-support automation includes rule-based chatbots, virtual assistants using wider information, tracking and return functions, and recommendation tools. Sentiment-analysis systems may flag language patterns for priority or review, but inferred patterns are not verified emotion and do not replace judgement. The questionnaire lists capability only as an exposure category. Reviews also find that conversational agents differ in task, information access and design, limiting broad effect claims (Mariani, Hashemi and Wirtz, 2023; Nicolescu and Tudorache, 2022).

Alharbi et al. (2025) synthesize technological, individual, organizational and contextual drivers and barriers across AI-chatbot adoption studies. The review synthesises adoption drivers and barriers, but adoption evidence does not establish whether a UK shopper's problem was resolved. Adoption serves as context, while task-level customer outcomes remain the evaluative focus.

Review evidence proposes routine automation, including support beyond staffed hours (Uzoka, Cadet and Ojukwu, 2024). Bahaw et al. (2025) instead examines SME perceptions of generative-AI usability, not out-of-hours service outcomes. Agentic systems may execute linked actions (McKinsey & Company, 2025), creating potential convenience but greater consequences from unsuitable action. This is a design implication consistent with Jha et al.'s (2026) conceptual analysis, not a measured retail effect.

Human-like language can make low-risk exchanges engaging while encouraging assumptions beyond system competence. After failure, tone does not replace error recognition, retained context and an authorized remedy. Trust is consequently separated from satisfaction, and human escalation is treated as recovery capability (Adam, Wessel and Benlian, 2021; Ozuem et al., 2024).

2.3 Chatbots, Interaction Quality and Human Presence

Adam, Wessel and Benlian's (2021) randomised experiment links verbal anthropomorphic cues with social presence and request compliance, not disclosure or accuracy. Ding and Najaf (2024) associate interactivity and perceived humanness with e-commerce trust, while Liao and Sundar (2022) argue for reliance calibrated to capability. Communication style may also influence attachment and satisfaction (Xu, Zhang and Deng, 2022). An appealing interface is not evidence of competence.

Recovery outcomes differ by study: Yun and Park (2022) examine satisfaction, repurchase intention and word of mouth; Ozuem et al. (2024) connect chatbot failure and recovery with loyalty; and Rese and Witthohn (2025) test warmth, competence, humanness and satisfaction. These studies support post-failure recovery, but do not directly test human transfer or preserved conversation context. Those requirements are assessed separately through the survey.

Social presence and competence may reinforce each other, but neither proves the other. Retailers must test conversational clarity, factual performance, task authority and escape from automation separately. The Hybrid Model prevents an appealing interface from masking a weak service result (Ding and Najaf, 2024; Chen et al., 2022).

Identity cues raise another trust and inclusion issue. Lu and Zhang (2025) examine how chatbot gender, race, name and job title affect consumer trust and preference across product contexts. The observations suggest design choices can affect perceived credibility and inclusion, but preference evidence does not establish the fairness or effectiveness of deployed retail support. The Hybrid Model treats anthropomorphic design as a possible trust cue requiring governance and task-level outcome evidence.

Expectation confirmation may explain post-use satisfaction, but a cross-sectional answer does not reveal whether the benchmark was a brand promise, earlier human service or task seriousness. The same reply may suit tracking and fail for a refund, so coefficients from other sectors are not directly transferable to UK retail (Nguyen, Chiu and Le, 2021; Chen et al., 2022; Ozuem et al., 2024).

2.4 AI Support, Satisfaction and Trust

Satisfaction is a post-use judgement, not an automatic AI effect. Nguyen and Le (2025) examine chatbot experience in banking; Zaheer, Sadiq and Safi (2025) examine service quality, trust and satisfaction across services; and Chau et al. (2025) studies e-commerce interaction and decisions. These cross-sectional, non-UK studies provide comparisons, not causal effects or UK population estimates.

Because much evidence is self-reported, cross-sectional and non-UK, direction and transfer cannot be assumed. Trust, satisfaction and task ease may share unmeasured causes; the investigation therefore reports full distributions and treats selected associations as exploratory (Bell, Bryman and Harley, 2022; Saunders, Lewis and Thornhill, 2023; Nguyen and Le, 2025).

Kumari, Gupta and Aeron (2025) offer a complementary generative-AI perspective. Their questionnaire study connects customers' perceptions of data capture and anthropomorphic interaction with trust, and considers how customer fatigue changes the expression of those relationships in citizenship behavior. This helps explain why a favorable interaction can coexist with hesitation or disengagement. Its self-selected international sample and behavioral intention focus nevertheless limit transfer to completed UK support encounters, so it refines interpretation here without validating a causal pathway.

Prior expectations change the meaning of ratings. A brief tracking update may exceed expectations, while the same automation may be unacceptable when money must be refunded or an account corrected. Satisfaction can thus reflect task difficulty and consequence as well as technology quality. Because the questionnaire did not capture issue severity, future operational use needs to segment overall ratings by task (Nguyen, Chiu and Le, 2021; Chen et al., 2022; Ozuem et al., 2024).

Outcome definitions help explain inconsistent observed AI effects. Responsiveness, purchase intention and continued use can appear more favorable than failure, repeat contact or completed redress, and routine enquiries have clearer endpoints than disputed transactions. AI-support performance is consequently conditional on task, authority, measure and available recovery (Chen et al., 2022; Ozuem et al., 2024; Rese and Witthohn, 2025).

Trust in AI accuracy is not the inverse of wanting human help. Q11 concerns confidence in an automated answer; Q16 concerns access to a person after failure. Both can be high because human access offers review and accountability. The survey examines coexistence, not a human-versus-AI comparison. Liao and Sundar (2022) supports capability-conditioned reliance; Rese and Witthohn (2025) is used only for recovery outcomes.

2.5 Service Quality, Speed and Resolution

Service quality is the mechanism through which technical capability may become customer value. AI-chatbot research separates responsiveness from reliability, assurance and problem-solving competence (Chen et al., 2022). First-response time is a readily observable operational metric, but it can favour an instant answer even when the underlying issue remains unresolved. Resolution and repeat contact are needed to determine whether the apparent efficiency benefited the customer.

For payments, returns and complaints, incomplete automation creates further contact and effort. Q5 measures perceived speed and Q8 perceived task effectiveness. A positive relationship would fit a functioning journey; a weak pattern would require separate ownership of the outcomes (Chen et al., 2022; Ozuem et al., 2024; Rese and Witthohn, 2025).

Resolution may mean accepted information, transaction correction, no repeat contact or a human remedy. For this dissertation, completion, reopening and repeat contact are proposed as retailer-side checks alongside customer judgement; they were not measured by the questionnaire. Chen et al. (2022) supports separating responsiveness, reliability and problem solving, while recovery studies demonstrate that post-failure outcomes require separate evaluation (Ozuem et al., 2024; Rese and Witthohn, 2025). The survey therefore captures perceived effectiveness, not verified operational resolution.

2.6 Personalization and Automation

Personalization can lower search effort by adapting information or recommendations to context. Research links usefulness, ease and suitable communication with experience and purchase intention (Lopes, Silva and Massano-Cardoso, 2024; Meng, Lu and Xu, 2025; Sidlauskienė, Joye and Auruskeviciene, 2023). However, relevance commonly relies on behavioral, preference or account data. Convenience and privacy can consequently improve or deteriorate at the same time.

As an author-derived governance recommendation, proportionate personalisation means limiting data to the task, explaining its use and retaining a less personalised or human route. Hasal et al. (2021) identifies chatbot privacy, security and data-protection risks, while Liao and Sundar (2022) supports communication that calibrates trust. Neither interface quality nor personalisation alone proves accuracy or accountability; trust and privacy therefore remain separate Hybrid Model dimensions.

Interface impressions must be distinguished from justified reliance. Human-like language may alter social presence or trust (Adam, Wessel and Benlian, 2021; Ding and Najaf, 2024), but it does not demonstrate factual accuracy or accountable data use. Direction is also uncertain in a one-time survey: a successful outcome may raise trust without trust necessarily producing it. This investigation therefore treats trust and privacy associations as exploratory and derives governance recommendations from the combined survey pattern, privacy evidence and responsible-trust literature rather than from one coefficient.

2.7 Privacy, Ethics and Responsible AI

Trust includes competence and conduct: the answer must be credible, and the organization's data practice must be acceptable. Transparency, interactivity and perceived humanness can influence chatbot trust, while excessive human resemblance may produce unjustified reliance (Alagarsamy and Mehroliya, 2023; Ding and Najaf, 2024; Liao and Sundar, 2022). Responsible design accordingly aims for calibrated trust, matched to a bounded task and supported by correction and review.

Conversational data create collection, retention, security and secondary-use risks (Hasal et al., 2021). UK retailers processing personal data are subject to the Data Protection Act 2018 (Data Protection Act 2018); this does not by itself demonstrate shopper understanding. Plain-language explanation, proportionate use and challenge routes are author-derived recommendations informed by privacy and responsible-trust literature (Hasal et al., 2021; Liao and Sundar, 2022).

Responsible trust means matching reliance to task evidence, authority and consequence (Liao and Sundar, 2022). In this dissertation, visible scope, uncertainty, correction and review are governance recommendations intended to keep accountability with the retailer rather than shifting error detection to the customer.

Digital inclusion shapes what evidence can be collected. An online questionnaire suits e-commerce users, but participation depends on access, confidence and willingness to use a digital form. Age alone does not represent exclusion: disability, language, affordability, literacy and temporary vulnerability can affect any group. Public reports give wider context; the survey records only those reached. Ease ratings remain conditional on participation and do not represent people absent from the dataset (Mubarak and Suomi, 2022; Lloyds Banking Group, 2025; Ofcom, 2026).

2.8 Digital Inclusion and Vulnerable Customers

Automation can disadvantage customers without reliable access, suitable devices, confidence or an accessible interface. Exclusion is not confined to age; income, disability, language, geography and temporary vulnerability may also matter (Mubarak and Suomi, 2022; Lloyds Banking Group, 2025). Ofcom documents continuing differences in UK confidence and participation (Ofcom, 2026). Consequently, an online survey is likely to under-observe some of the people most affected.

AI agents may simplify information for vulnerable customers, yet opacity or an inflexible journey can deepen the same vulnerability (Jha et al., 2026). Q13 captures perceived ease, not a complete accessibility assessment. Retailers still need direct testing across devices, abilities and language needs, plus an alternative channel when the digital route cannot be used.

Accessibility extends beyond opening a page: reading complexity, contrast, authentication, devices and confidence to challenge an answer affect completion. Conformance can still leave excessive cognitive effort or unusable recovery. Q13 represents a broad ease perception, not an accessibility audit; inclusive performance requires direct testing across access needs (Mubarak and Suomi, 2022; Ofcom, 2026; Jha et al., 2026).

For this dissertation, the author classifies UK sources by evidential role. ONS describes retail activity; Ofcom and Lloyds illuminate participation and capability; and industry material provides timely organizational context. These sources do not measure the same construct and are not pooled. Official and public evidence establishes the operating environment, while peer-reviewed research supplies mechanisms and limitations against which the survey is interpreted (ONS, 2026a; Ofcom, 2026).

2.9 UK Online Retail Context

UK online retail combines substantial transaction volume with high-consequence after-sales tasks. ONS data establish the importance of internet sales (ONS, 2026a; ONS, 2026b), and public evidence shows the harm created by poor parcel resolution (Citizens Advice, 2025; Ofcom, 2025). Automation may be proportionate for product information or tracking, whereas disputed payments, repeated failure and vulnerable customers require stronger review and recovery.

ONS (2026c) sector-level internet-retail data show that online retail is not homogeneous. These figures provide context but do not identify AI use or customer outcomes. Future research needs to record retailer and support task instead of generalizing one journey-stage distribution across all retail activity.

Wadud, Liang and Polkinghorne (2025) situate UK retail digital transformation in changes to both operational processes and customer interaction. This supports treating customer service as part of retail transformation and not as a purely technical deployment. A broad sector account does not, however, establish whether a particular chatbot is trustworthy, inclusive or effective. Those questions require customer-level and task-level evidence, which the Hybrid Model keeps analytically distinct.

Technology-acceptance research addresses usefulness and effort (Auer, Schlögl and Glowka, 2024; Lopes, Silva and Massano-Cardoso, 2024); AI-chatbot service-quality research separates delivered performance dimensions (Chen et al., 2022); expectation confirmation concerns post-use judgement (Nguyen, Chiu and Le, 2021); and responsible-trust theory addresses reliance under uncertainty (Liao and Sundar, 2022). Combining them into one score would risk circular explanation. The Hybrid Model therefore preserves separate constructs: weak relationships do not disprove each lens, but challenge a simple integrated pathway in this sample.

2.10 Theoretical Frameworks

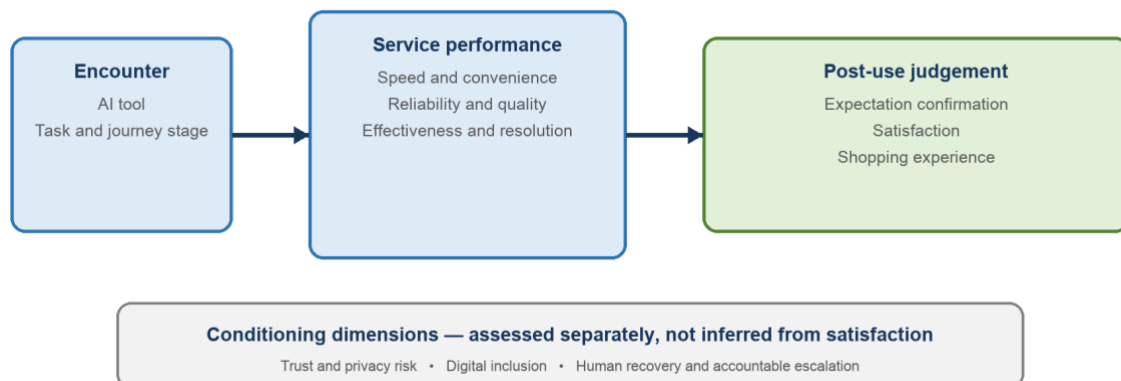
Table 2.1 Theoretical lenses and analytical use

Lens	Analytical focus	Supporting sources	Use in this study
Service-quality lens	Responsiveness, reliability, competence and resolution are examined separately.	Chen et al. (2022); Chau et al. (2025)	Prevents speed from standing for the whole service.
Technology-acceptance lens	Usefulness and ease help interpret adoption and continuance.	Auer, Schlögl and Glowka (2024); Lopes, Silva and Massano-Cardoso (2024)	Applied cautiously because adoption is not completed support.
Responsible-trust and privacy lens	Trust should be calibrated to capability and data practice.	Liao and Sundar (2022); Hasal et al. (2021)	Separates confidence from blind reliance.
Digital-inclusion and service-recovery lens	Access, confidence and vulnerability are distinguished from recovery after chatbot failure.	Mubarak and Suomi (2022); Jha et al. (2026); Ozuem et al. (2024); Rese and Witthohn (2025)	Uses inclusion sources for participation and recovery sources for post-failure satisfaction or loyalty.

Source: Author synthesis of the retained 40-source evidence corpus.

Figure 2.1 Combined evidence framework for AI customer support

Hybrid diagnostic framework for UK online retail AI support



Source: Author developed analytical framework based on the reviewed literature; it is not a validated scale or causal model.

The four lenses address different stages of customer judgement and therefore cannot be treated as substitutes. Service-quality theory evaluates delivered responsiveness, reliability, assurance and empathy, whereas technology-acceptance research explains willingness to use a channel through perceived usefulness and effort. Trust-risk theory asks whether reliance is justified despite accuracy, performance and privacy risks, while expectation confirmation evaluates the gap between prior expectations and experienced performance (Chen et al., 2022; Auer, Schlögl and Glowka, 2024; Alagarsamy and Mehroli, 2023; Nguyen, Chiu and Le, 2021).

Their assumptions also create productive tensions. An interface may be easy and useful yet unreliable, fast yet unsatisfactory, or satisfactory in one low-risk task while inappropriate for a consequential return or payment dispute. Acceptance consequently explains channel use but not service success; service quality assesses delivery but not necessarily justified trust; expectation confirmation depends on prior standards that may differ between customers. Keeping the constructs separate prevents a favorable rating on one dimension from becoming evidence for the others (Chen et al., 2022; Liao and Sundar, 2022).

Each lens has a transfer boundary in UK online retail. AI-specific service-quality measures remain sensitive to tool and task design; TAM evidence from another service sector cannot establish retail outcomes; trust-risk relationships may change with data sensitivity and system authority; and expectation confirmation does not identify whether the technology caused the experience. The theories also operate at different units of analysis: acceptance concerns an individual's willingness to use, service quality concerns an evaluated encounter, and responsible reliance concerns whether a system should be trusted for a particular task. Combining them without preserving those units would create conceptual overlap and overstate what one cross-sectional questionnaire can test. The lenses are therefore used as diagnostic perspectives, not as a validated composite scale or causal chain (Chen et al., 2022; Nguyen, Chiu and Le, 2021; Liao and Sundar, 2022).

2.11 The Hybrid Model for UK Online Retail

No single lens covers the revised objectives. The Hybrid Model links tool and task to four analytical expectations: responsiveness need not imply reliability or satisfaction; trust and privacy risk remain distinct from satisfaction; expectation confirmation should overlap only partly with post-use satisfaction; and digital confidence alone cannot represent inclusive access. Human escalation is treated as a recovery condition when automated capability or authority is insufficient (Chen et al., 2022; Liao and Sundar, 2022; Mubarak and Suomi, 2022).

These expectations organize rather than predetermine the analysis. Chapter 4 examines whether survey distributions and pre-specified associations support, challenge or leave each expectation unresolved, then compares that judgement with the retained literature. Support requires convergence between the direction of the survey pattern and a credible mechanism in the reviewed evidence; contradiction is retained where the observed pattern runs against the expectation; an expectation remains unresolved when imprecision, measurement limits or rival explanations prevent a defensible judgement. Contradictory and negligible relationships remain analytically important because they identify where the model should not be simplified or interpreted causally (Chen et al., 2022; Ozuem et al., 2024).

The retained corpus did not provide directly comparable UK customer evidence spanning tools, service outcomes, satisfaction, trust, privacy, inclusion and recovery in one design. The study addresses that bounded gap by applying the diagnostic framework to 147 reported encounters and triangulating the resulting patterns with the reviewed evidence. It does not estimate retailer prevalence, verify deployments or validate causal pathways (Mariani, Hashemi and Wirtz, 2023; Alharbi et al., 2025; Wadud, Liang and Polkinghorne, 2025).

2.12 Research Gap and Chapter Summary

The literature reveals a fragmented evidence base: capability, responsiveness, satisfaction, privacy, exclusion and recovery are commonly examined with different systems, sectors and outcomes. The contribution is therefore not a universal AI-effect estimate, but a UK online retail application of a diagnostic, non-causal framework that keeps those dimensions comparable without collapsing them into one score (Mariani, Hashemi and Wirtz, 2023; Chen et al., 2022).

The evidence hierarchy constrains that contribution. Peer-reviewed chatbot and e-commerce studies offer constructs and plausible mechanisms, but many use non-UK, cross-sectional or single-system samples. Official sources establish UK conditions without estimating customer-level relationships, while industry evidence is current but may privilege adoption narratives. A survey-literature match is therefore interpreted as corroboration of practical relevance, not replication: differences in population, task, system authority and outcome definition remain. Conclusions require convergence across source types and explicit treatment of disagreement and transfer limits (Chen et al., 2022; Chau et al., 2025; Ding and Najaf, 2024).

Accordingly, Chapter 4 evaluates four theoretical expectations against complete survey distributions, uncertainty intervals and the retained evidence. Support is graded as convergent, challenged or unresolved; none of these judgements validates a composite model or establishes cause. This approach also guards against adoption-led reasoning: the presence or frequency of an AI tool is not treated as evidence of service quality, customer benefit or responsible operation. The research gap is addressed only to the extent permitted by the achieved sample and the transferability of the retained sources (Chen et al., 2022; Liao and Sundar, 2022).

Chapter 3: Research Methods and Methodology

3.1 Introduction

This chapter explains the combined design used for the five revised objectives. Primary evidence came from 147 eligible survey respondents and secondary evidence from the retained 40-source evidence corpus. The procedure identifies the AI tools, describes performance and satisfaction, explores advantages and constraints, interprets experience and develops recommendations without estimating national prevalence or causality.

3.2 Research Philosophy and Approach

This investigation adopts a post-positivist position. Ratings are treated as imperfect indicators of experience, not direct technology effects or retailer-deployment records. The approach is deductive because the literature and revised objectives specified tools, performance, satisfaction, trust, privacy, exclusion, expectations and recovery before interpretation. Weak and unexpected findings were retained (Bell, Bryman and Harley, 2022).

This position is consistent with business-research guidance that methods should fit the question, make assumptions visible and report limits to inference (Bell, Bryman and Harley, 2022; Saunders, Lewis and Thornhill, 2023). It supports quantitative coding and association analysis while recognizing that cross-sectional self-report data cannot establish temporal order, mechanism or cause.

3.3 Combined Research Design

The primary component is a cross-sectional quantitative survey because RQ1-RQ4 require a consistent description of recent tools, tasks and ordered customer evaluations across the same encounters. Asking each eligible participant about the most recent AI support experience reduced ambiguity from averaging across retailers, while fixed response categories enabled transparent distributions and ordinal associations. An experimental design would better identify effects but would require control over retailer systems and exposure; qualitative interviews would reveal mechanisms but would not provide the same standardized distributions across 147 encounters. The selected design therefore fits the descriptive and exploratory questions, while its inability to establish temporal order, mechanism or technology effects is explicit (Bell, Bryman and Harley, 2022).

The structured critical review complements rather than repairs the survey. It is appropriate to RQ5 because responsible recommendations require evidence beyond respondents' preferences, including service quality, trust, privacy, inclusion and recovery research. Triangulation tests whether the primary pattern converges with, diverges from or remains unresolved by the retained 40-source corpus. It does not convert the project into a meta-analysis or make the convenience sample representative (Bell, Bryman and Harley, 2022).

3.4 Population, Sampling and Achieved Sample

The target population was UK residents aged 18 or over who had shopped online and encountered AI support during the preceding 12 months. The practically accessible population comprised eligible users reached through Discord, Facebook and X between 3 and 31 July 2026. From that accessible population, voluntary-response convenience sampling produced the achieved sample of 147 complete eligible respondents. This distinction matters: the sample is relevant to the research questions, but members of the target population had neither a known nor equal selection probability.

Eligibility was enforced through consent, age/residence and recent-exposure screening; all 147 supplied records were eligible and complete. Open recruitment exceeded the approximately 70 respondents anticipated in the ethics form without changing recruitment or eligibility. Openings, refusals and abandonments were not retained, so no response rate can be calculated, and no a priori power calculation was conducted. Sample adequacy is judged from achieved precision: Wilson intervals for overall top-two proportions are informative but remain several percentage points wide, while subgroup and correlation intervals are materially wider. Complete-case availability removes item-level denominator loss except for the declared Prefer not to say responses, but it does not reduce coverage or self-selection error. Precision is consequently reported for each selected estimate instead of applying a retrospective pass/fail sample-size claim. The sample supports bounded overall distributions and exploratory pooled comparisons, not national prevalence or precise tool-task-age estimates (Bell, Bryman and Harley, 2022).

3.5 Questionnaire and Measures

The approved instrument contained three screening, four profile and 19 substantive questions, all unchanged. Objective 1 uses Q1, with Q2-Q4 providing context. Objective 2 uses Q5-Q10. Objective 3 draws on P3-P4 and Q4, Q6-Q14. Objective 4 uses Q3, Q10 and Q15-Q18 to interpret stage, satisfaction, expectations, resolution and human access. Objective 5 uses Q19 and triangulation across all findings.

Single items reduced respondent burden and preserved direct alignment with the approved 19 substantive questions, but they do not form validated multi-item constructs. Table 3.1 and Appendix B map speed, reliability and quality to service performance; convenience and ease to enabling conditions; satisfaction and met expectations to post-use judgement; trust and privacy to justified reliance; and confidence, barriers and human access to inclusion and recovery. The operationalization deliberately keeps conceptually related indicators separate, allowing a fast but ineffective encounter or an easy but untrusted interaction to remain visible. The trade-off is reduced depth: one response cannot capture all facets of reliability, trust, privacy or accessibility, and differences between indicators may include wording effects. This provides face and content alignment, not construct validation. Internal-consistency coefficients are inapplicable to separate single-item indicators, and no retrospective pilot claim is made because no pilot or cognitive-interview report is retained (Bell, Bryman and Harley, 2022).

Table 3.1 Objective, question and evidence mapping

Objective	RQ	Survey measures	Secondary-review themes
1	RQ1	Q1 primary; Q2-Q4: task, stage and relevance context	AI tools and customer-support uses in online retail
2	RQ2	Q5-Q10: performance dimensions and satisfaction	Service performance, quality and customer satisfaction
3	RQ3	P3-P4; Q4 and Q6-Q14	Benefits, limitations, trust, privacy and digital exclusion
4	RQ4	Q3, Q10 and Q15-Q18	Online-shopping experience, expectations, resolution and human recovery
5	RQ5	Q19 plus triangulation	Responsible and effective retailer use

Source: Author revised mapping using the unchanged approved questionnaire and retained 40-source evidence corpus.

3.6 Secondary Research Procedure

Secondary investigation was conducted during June and July 2026 through Google Scholar, UWTSD library resources, publisher pages and official UK websites. The documented query logic combined ("AI chatbot" OR "conversational agent" OR "AI customer service") with ("online retail" OR e-commerce) and one or more outcome blocks: service quality/satisfaction, trust/privacy, accessibility/digital inclusion, or recovery/human escalation. UK-context searches additionally combined online retail with ONS, Ofcom, parcel delivery, digital participation and consumer redress. Searches prioritized peer-reviewed research and official UK evidence; industry material was retained only for current organizational capability or adoption context.

The retained 40-source evidence corpus was assessed against five explicit criteria: relevance to at least one objective; identifiable study design or source authority; sector and geographical fit; ability to support the specific claim type; and a stated transfer limitation where the evidence was non-UK, non-retail, cross-sectional, conceptual or industry-produced. Appendix E records method/design, sector/geography, evidential role and transfer limitation for every source. A database-by-database result log and historical exclusion counts were not retained, so none is reconstructed. The procedure is therefore reproducible at the level of query logic and appraisal rules, but remains a structured bounded review rather than an exhaustive systematic review or meta-analysis (Bell, Bryman and Harley, 2022).

3.7 Data Preparation, Coding and Quality Control

The original Answers worksheet was preserved. A separate coded worksheet linked each record to visible coding rules. Ordered five-category responses were coded from 1 to 5 in questionnaire order. For Q12, a higher score explicitly means greater privacy concern. Q14 was coded from 1 for No to 4 for Yes, completely; seven Prefer not to say responses were unavailable for ordinal analysis, giving valid $n=140$. Any reported barrier was derived as 1 when at least one named P4 barrier was marked Yes and 0 otherwise.

Quality checks confirmed 147 unique record IDs, complete responses, valid categories, no duplicated complete answer patterns and no conflicting selection of No barriers with a named barrier. The codebook, formula-linked coded data, frequency tables and dissertation tables provide an auditable path from raw answers to reported results. Percentages are calculated against the stated valid denominator, and rounding is to one decimal place.

Table 3.2 Coding and analysis rules

Variable(s)	Coding	Primary use	Boundary
Q4–Q13, Q15–Q16	Ordinal 1–5 in administered order	Full distribution, median, top-two rate	No composite score
Q12	1=not at all concerned; 5=extremely concerned	Higher score=greater privacy concern	Direction checked separately
Q14	1=No; 2=Only partly; 3=Yes, mostly; 4=Yes, completely	Ordinal valid $n=140$	Prefer not to say unavailable
P3	1=not at all to 5=extremely confident	High=4–5; lower=1–3	Descriptive groups only
P4	Any named barrier Yes=1; otherwise 0	Barrier $n=25$; no named barrier $n=122$	Does not measure every form of exclusion

Source: Author coding protocol; full codebook supplied in Survey-Analysis.xlsx.

3.8 Analysis and Triangulation Strategy

Counts, percentages, complete distributions and medians are primary because the measures are categorical or ordinal. Wilson 95% intervals accompany selected top-two proportions to expose uncertainty rather than imply representativeness. Low-frequency tool categories were pooled into Other/pooled ($n=28$), and tasks into four service-function groups, to avoid unstable single-digit reporting. Pooling improves descriptive stability but does not make unlike technologies or task severities equivalent; all group contrasts remain exploratory (Bell, Bryman and Harley, 2022).

Spearman correlation was selected because the eleven pairs combine ordered ratings and require a monotonic, rank-based measure without assuming interval spacing or normality. The pairs were pre-specified from the Hybrid Model rather than selected after examining every possible association. Percentile 95% intervals from 5,000 deterministic bootstrap resamples show coefficient uncertainty and the stability of direction under resampling of the achieved records. Because eleven relationships are inspected, isolated non-zero intervals are not treated as confirmatory discoveries; interpretation emphasizes magnitude, interval width, coherence and contradiction across the framework. No p-value search, multiplicity-based selection, regression, factor analysis or composite scale was used; the coefficients are descriptive evidence within the achieved sample, not causal effects or confirmatory population tests (Bell, Bryman and Harley, 2022).

Triangulation followed the five objectives and used an explicit evidence hierarchy. For each main survey pattern, the analysis identified the closest relevant sources, compared direction and construct meaning, and then recorded convergence, disagreement or an unresolved relationship. Peer-reviewed empirical evidence received greatest weight for mechanisms; official UK sources established context; industry evidence informed current practice only where its promotional incentives were acknowledged. Convergence increased the practical plausibility of an interpretation but did not repair convenience sampling, common-method measurement or cross sector transfer. Divergence was examined for rival explanations such as task mix, customer expectations and measurement scope rather than removed as noise. Recommendations were retained only when the survey, literature or both supplied an identifiable rationale, and their evidential basis is stated in Table 5.1 (Bell, Bryman and Harley, 2022).

3.9 Reliability, Validity and Ethics

Procedural reliability derives from fixed wording, preserved answers, explicit coding, formula-linked outputs and numerical reconciliation. Reproducibility applies to the transformation of the supplied records into reported tables, not to the stability of an individual's rating across time, which was not retested. Content validity is supported by the objective-question mapping and literature-derived dimensions, while face validity concerns whether the administered items appear relevant to the stated constructs. Construct validity remains limited because most constructs use one item and different AI technologies are grouped internal consistency cannot be estimated for independent single items. External validity is weakest because voluntary online recruitment does not support population generalization. Triangulation improves interpretive discipline by confronting findings with different evidence, but it cannot compensate for measurement or selection bias. The framework therefore supports transparent, reproducible description rather than scale validation (Bell, Bryman and Harley, 2022).

The final ethics form records supervisor sign-off by Dr Lakshani Mendis on 25 April 2026, before the 3–31 July data collection. It classifies the online questionnaire with UK adults aged 18 or over as low risk. Participation was voluntary and based on informed consent; respondents could omit questions or stop before submission. The instrument prohibited identifying, account and payment information. Analysis used anonymous record IDs and aggregate reporting. Possible mild discomfort concerning trust, privacy or negative support experiences, and limited data-protection risk, were mitigated through data minimization, confidentiality and restricted password-protected storage. The form contains no committee name or approval reference, so the approval is described as supervisor sign-off. Appendix F records the protocol (Bell, Bryman and Harley, 2022).

3.10 Methodological Limitations

Convenience recruitment creates unknown coverage and self-selection error, particularly because an online survey may underrepresent people facing the greatest digital barriers. Recall, interpretation and common-method effects may also influence associations between ratings from the same encounter. The achieved sample is insufficient for precise original tool/task/age barrier comparisons, and the most recent encounter may not represent a respondent's typical experience (Bell, Bryman and Harley, 2022; Saunders, Lewis and Thornhill, 2023; Mubarak and Suomi, 2022).

The design cannot determine whether AI caused performance, satisfaction or experience changes, verify retailer system classifications, or equate reported completion with operational resolution. P4 is not a complete accessibility assessment, and differences across the retained literature limit transfer. These constraints are carried into every interpretation and recommendation (Bell, Bryman and Harley, 2022).

3.11 Chapter Summary

The design transparently links 147 eligible responses and the retained evidence corpus to the revised objectives. Its strengths are traceability and auditable outputs; its limits are non-probability sampling, single-item measurement, self-report and unverified retailer-system data. Chapter 4 reports the observations by objective.

Chapter 4: Findings, Analysis and Discussion

4.1 Data Quality and Sample Profile

All 147 records passed the consent, UK-resident age-18-plus and recent AI-support-exposure screens. Every analyzed field was complete, record IDs were unique, response categories matched the administered questionnaire and no complete answer pattern was duplicated. The quality audit therefore found no record-level exclusion after eligibility screening. This completeness supports internal reconciliation, although it does not remove selection or recall bias.

Age was distributed across six groups: 25 respondents were aged 18–24, 35 were 25–34, 30 were 35–44, 25 were 45–54, 20 were 55–64 and 12 were 65 or over. Eighty-three respondents (56.5%) described themselves as very or extremely digitally confident, while 64 (43.5%) selected the lower three categories. Twenty-five respondents (17.0%) reported at least one named participation barrier. The most frequent named barriers were security or privacy concern (11), limited digital skills (9), needing help (8) and limited internet or device access (5). These counts show diversity within the sample but cannot indicate UK prevalence.

Table 4.1 Data quality and sample profile

Indicator	Count	Percent	Definition / note
Eligible complete records	147	100.0%	All screens passed; no missing analyzed values
Age 18–24	25	17.0%	Age distribution
Age 25–34	35	23.8%	Largest age group
Age 35–44	30	20.4%	Age distribution
Age 45–54	25	17.0%	Age distribution
Age 55–64	20	13.6%	Age distribution
Age 65 or over	12	8.2%	Smallest age group
High digital confidence	83	56.5%	Very or extremely confident
Lower digital confidence	64	43.5%	Not at all, slightly or moderately confident
Any named barrier	25	17.0%	At least one named P4 barrier
No named barrier	122	83.0%	No named P4 barrier reported

Source: Author analysis of the 147 eligible survey responses.

4.2 Main AI Tools in UK Online Retail Customer Support

Text-based chatbots were the leading identified tool (80/147; 54.4%), followed by tracking or delivery support (23) and AI help-center or search tools (16). The remaining categories were individually uncommon. These customer-reported encounters describe the achieved sample, not retailer market share or verified system deployment.

Order or delivery checks and returns or exchanges were the joint-leading tasks (32 each), while return, exchange or refund was the largest journey stage (45). The distribution places much reported AI support after purchase, where access to correction, money or redress may increase the consequence of inaccurate or unauthorized action.

Relevance was positive for 95 respondents (64.6%; median 4), but tool-by-task counts show that one tool category served materially different purposes. Chatbots, for example, covered delivery checks, returns and other problems. Tool identity alone is therefore an incomplete unit of evaluation; task and journey stage provide necessary context.

Table 4.2 Leading tools, tasks and journey stages

Type	Category	Count	Percent
Tool	Text-based chatbot	80	54.4%
Tool	Automated order-tracking or delivery support	23	15.6%
Tool	AI-powered help center or search	16	10.9%
Task	Checking an order or delivery	32	21.8%
Task	Returning or exchanging a product	32	21.8%

Task	Finding product information	21	14.3%
Stage	Return, exchange or refund	45	30.6%
Stage	Before purchase	27	18.4%
Stage	During purchase or payment	27	18.4%
Relevance	Very or extremely relevant	95	64.6%

Source: Author analysis of Q1–Q4 (n=147).

Reported AI-enabled support tools

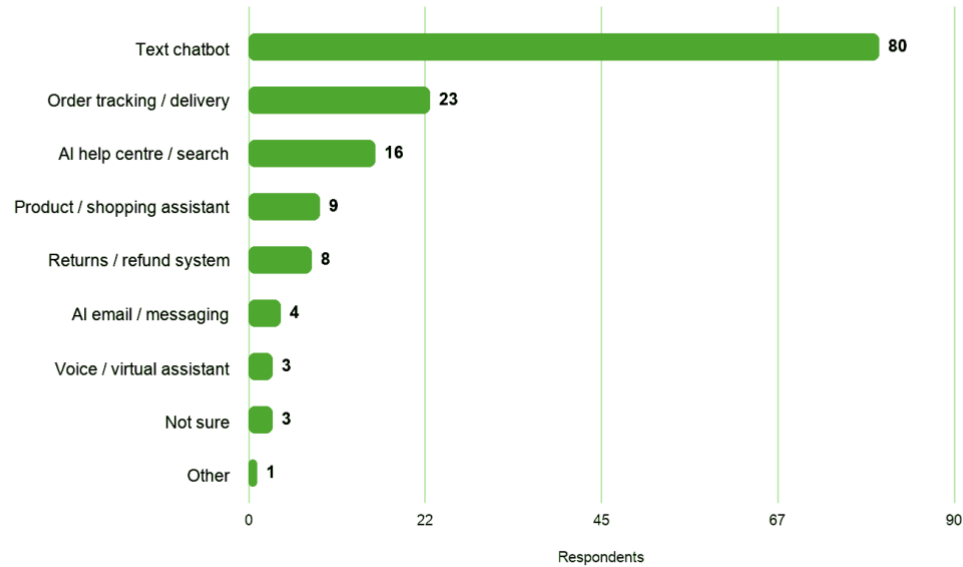


Figure 4.1 Reported AI support tools

Source: Author analysis of Q1 (n=147).

Main customer-service tasks

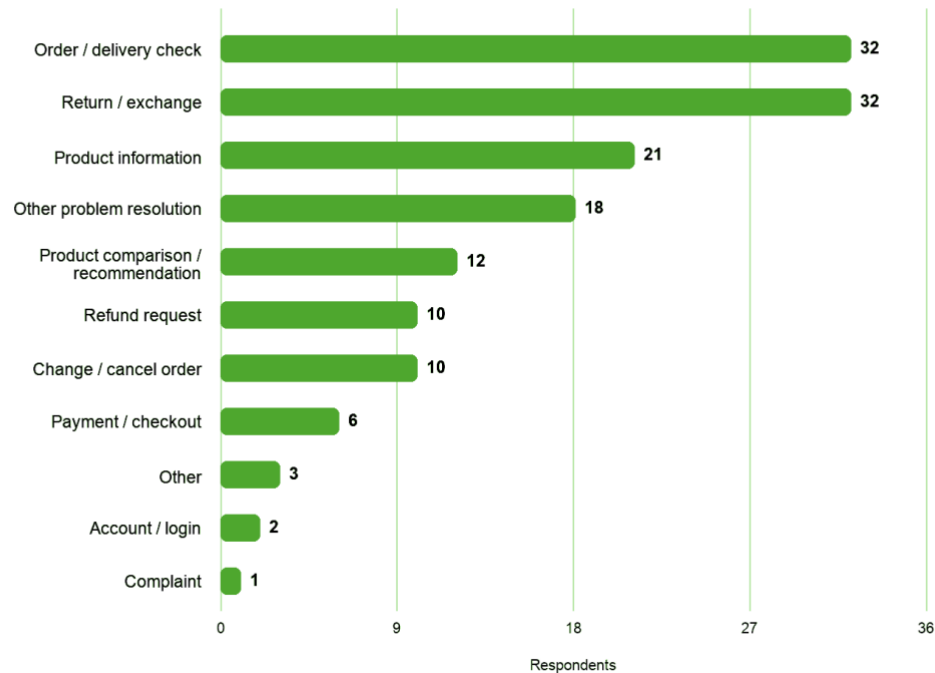


Figure 4.2 Main customer-service tasks

Source: Author analysis of Q2 (n=147).

The secondary evidence supports a capability-based interpretation of retailer use. Reviews describe conversational agents, recommendations and automated service functions as different systems with different information access and authority (Mariani, Hashemi and Wirtz, 2023; Nicolescu and Tudorache, 2022). It supports the tool taxonomy but cannot establish UK retailer prevalence. The survey adds customer-reported evidence of substantial post-purchase encounters, where correction, money and redress can raise the consequence of failure.

4.3 Service Performance and Customer Satisfaction

Speed was the strongest-rated performance dimension (71.4%; Wilson 95% CI 63.7-78.1), followed by convenience (68.7%). Reliability, service quality, effectiveness and satisfaction were less favorable, ranging from 55.1% to 59.2% positive. Although all medians were 4, the complete distributions and intervals show that a quick response was not equivalent to dependable resolution or satisfaction.

The gaps between speed and effectiveness, service quality and satisfaction were 15.6, 12.2 and 16.3 percentage points. They are comparisons between judgements from the same encounters, not treatment effects, and unmeasured task difficulty or prior expectations may contribute. Retailers should therefore avoid using response time as a proxy for completed service (Chen et al., 2022; Ozuem et al., 2024; Rese and Witthohn, 2025).

Table 4.3 Full support-performance distributions

Measure	1	2	3	4	5	Median
Response speed	3 (2.0%)	13 (8.8%)	26 (17.7%)	56 (38.1%)	49 (33.3%)	4
Reliability	8 (5.4%)	18 (12.2%)	35 (23.8%)	59 (40.1%)	27 (18.4%)	4
Convenience	13 (8.8%)	13 (8.8%)	20 (13.6%)	57 (38.8%)	44 (29.9%)	4
Effectiveness	8 (5.4%)	19 (12.9%)	38 (25.9%)	62 (42.2%)	20 (13.6%)	4
Service quality	9 (6.1%)	13 (8.8%)	38 (25.9%)	63 (42.9%)	24 (16.3%)	4

Source: Author analysis of Q5–Q9 (n=147). Cells report count (percent). Scores run from the least favorable category (1) to the most favorable category (5).

Positive support-performance ratings

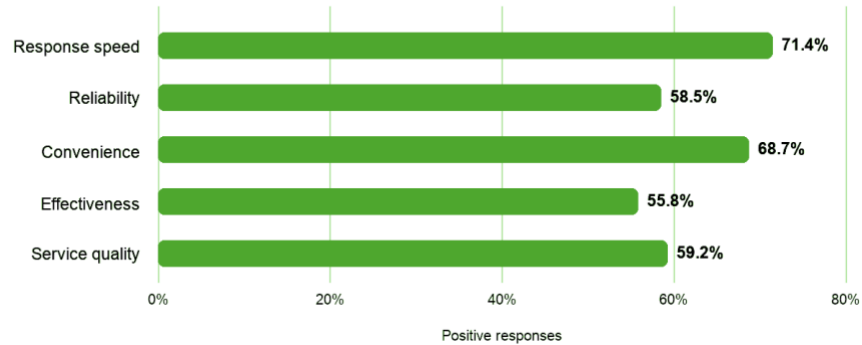


Figure 4.3 Positive support-performance ratings

Source: Author analysis of Q5–Q9 (n=147); positive means the two most favorable categories.

Exploratory comparisons were added to test whether the aggregate pattern concealed different tool or task profiles. Original categories were pooled only where necessary to avoid reporting single-digit tool cells or highly fragmented tasks. Each cell below reports the positive top-two percentage and its Wilson 95% confidence interval; no between-group significance test or causal ranking is claimed (Bell, Bryman and Harley, 2022).

Table 4.4 Exploratory tool-group performance and satisfaction

Group (n)	Speed	Reliability	Convenience	Effectiveness	Service quality	Satisfaction
Text chatbot (n=80)	67.5% (56.6-76.8)	63.8% (52.8-73.4)	68.8% (57.9-77.8)	56.3% (45.3-66.6)	63.8% (52.8-73.4)	51.3% (40.5-61.9)
Tracking/delivery (n=23)	78.3% (58.1-90.3)	69.6% (49.1-84.4)	60.9% (40.8-77.8)	60.9% (40.8-77.8)	39.1% (22.2-59.2)	73.9% (53.5-87.5)
Help center/search (n=16)	68.8% (44.4-85.8)	50.0% (28.0-72.0)	75.0% (50.5-89.8)	56.3% (33.2-76.9)	62.5% (38.6-81.5)	68.8% (44.4-85.8)
Other/pooled (n=28)	78.6% (60.5-89.8)	39.3% (23.6-57.6)	71.4% (52.9-84.7)	50.0% (32.6-67.4)	60.7% (42.4-76.4)	42.9% (26.5-60.9)

Source: Author analysis of Q1 and Q5-Q10 (n=147). Other/pooled combines low-frequency and uncertain tool categories; cells show top-two percent (Wilson 95% CI).

Table 4.5 Exploratory task-group performance and satisfaction

Group (n)	Speed	Reliability	Convenience	Effectiveness	Service quality	Satisfaction
Information/comparison (n=33)	78.8% (62.2-89.3)	51.5% (35.2-67.5)	75.8% (59.0-87.2)	54.5% (38.0-70.2)	54.5% (38.0-70.2)	51.5% (35.2-67.5)
Order/delivery check (n=32)	75.0% (57.9-86.7)	59.4% (42.3-74.5)	71.9% (54.6-84.4)	46.9% (30.9-63.6)	68.8% (51.4-82.0)	62.5% (45.3-77.1)
Return/refund/change (n=52)	63.5% (49.9-75.2)	63.5% (49.9-75.2)	63.5% (49.9-75.2)	67.3% (53.8-78.5)	51.9% (38.7-64.9)	53.8% (40.5-66.7)
Other service/problem (n=30)	73.3% (55.6-85.8)	56.7% (39.2-72.6)	66.7% (48.8-80.8)	46.7% (30.2-63.9)	66.7% (48.8-80.8)	53.3% (36.1-69.8)

Source: Author analysis of Q2 and Q5-Q10 (n=147). Tasks were pooled by service function; cells show top-two percent (Wilson 95% CI).

Exploratory profiles varied but did not establish a stable tool ranking. Tracking/delivery support combined comparatively high satisfaction with lower service quality, whereas returns/refunds tasks showed lower speed but higher effectiveness than order/delivery checks. Overlapping intervals, heterogeneous pooled categories and uncontrolled task seriousness offer competing explanations for those contrasts. They identify retailer-level testing questions rather than comparative effects (Mariani, Hashemi and Wirtz, 2023; Chen et al., 2022).

All selected associations were weak or negligible. Speed-convenience was largest ($\rho=.225$, 95% bootstrap CI .064 to .374), satisfaction-expectations was .207 (.046 to .356), and effectiveness-satisfaction was .131 (-.035 to .303). Wide intervals and six intervals crossing zero indicate that no single rating reliably predicts the broader experience within this design (Bell, Bryman and Harley, 2022).

This evidence partly supports service-quality theory's separation of responsiveness, reliability and problem solving. Alternative explanations remain: different tasks may set different performance thresholds, and single-item ratings may not capture multidimensional service quality. The practical judgement is therefore to monitor time, accuracy, resolution, repeat contact and satisfaction separately rather than calculate one aggregate AI score (Chen et al., 2022; Auer, Schlögl and Glowka, 2024; Chau et al., 2025).

4.4 Benefits and Limitations: Trust, Privacy and Digital Exclusion

Benefits were conditional rather than uniform. Relevance, convenience and ease were positive for roughly two-thirds of respondents, while satisfaction was positive for 55.1%, strong trust for 36.7% and high privacy concern for 36.1%. Convenience can therefore coexist with cautious reliance and concern about personal data; neither satisfaction nor ease demonstrates that trust is justified (Liao and Sundar, 2022; Hasal et al., 2021).

Independent completion was reported by 111 of 140 valid respondents, but it cannot establish accessibility or successful remedy. A user may finish an interaction without understanding the outcome, obtaining an authorised resolution or having an effective alternative. Completion, accessibility and service success must consequently remain separate indicators (Mubarak and Suomi, 2022; Jha et al., 2026; Chen et al., 2022).

Table 4.6 Satisfaction, trust, privacy, ease and expectation distributions

Measure	1	2	3	4	5	Median
---------	---	---	---	---	---	--------

Satisfaction	15 (10.2%)	21 (14.3%)	30 (20.4%)	51 (34.7%)	30 (20.4%)	4
Trust	11 (7.5%)	28 (19.0%)	54 (36.7%)	43 (29.3%)	11 (7.5%)	3
Privacy concern	23 (15.6%)	22 (15.0%)	49 (33.3%)	38 (25.9%)	15 (10.2%)	3
Ease of use	3 (2.0%)	13 (8.8%)	32 (21.8%)	61 (41.5%)	38 (25.9%)	4
Met expectations	10 (6.8%)	25 (17.0%)	32 (21.8%)	58 (39.5%)	22 (15.0%)	4
Human-adviser importance	1 (0.7%)	5 (3.4%)	19 (12.9%)	61 (41.5%)	61 (41.5%)	4

Source: Author analysis of Q10–Q13 and Q15–Q16 (n=147). Cells report count (percent). For privacy, a higher score means greater concern; for other rows, a higher score is more favorable or more important.

Satisfaction, trust, privacy, ease and completion

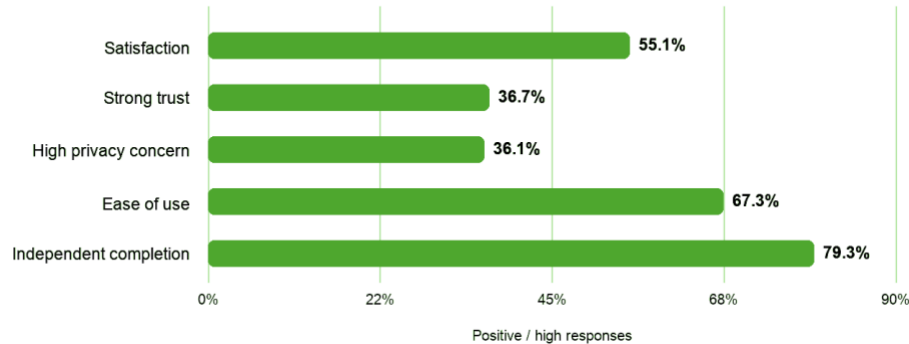


Figure 4.4 Satisfaction, trust, privacy, ease and independent completion

Source: Author analysis of Q10–Q14. Q14 independent completion uses valid n=140; other measures use n=147.

Trust-satisfaction ($\rho = -.022$) and trust-privacy concern (.026) were negligible; service quality-privacy concern was weakly negative (-.176). These patterns challenge any simple trust-risk pathway but do not refute the theory. Heterogeneous task sensitivity and single-item measurement could attenuate relationships, while restricted self-selection could omit customers who avoid AI support. Trust, privacy and satisfaction should therefore be assessed independently (Bell, Bryman and Harley, 2022; Liao and Sundar, 2022; Hasal et al., 2021).

Higher digital confidence did not produce consistently better evaluations: the high-confidence group reported lower positive convenience and ease, with weak or negligible negative correlations. Greater expectations among confident users and a different task mix are plausible alternatives, but neither was measured. Confidence alone is therefore an inadequate indicator of inclusive service (Mubarak and Suomi, 2022; Lloyds Banking Group, 2025; Ofcom, 2026).

The named-barrier group reported lower speed, reliability and strong trust but slightly higher satisfaction and effectiveness. This internally mixed pattern may reflect the small heterogeneous subgroup (n=25), self-selection into an online survey or uncontrolled task differences. It neither establishes nor dismisses disadvantage; it shows why retailers need direct accessibility and recovery testing rather than inference from average ease (Mubarak and Suomi, 2022; Jha et al., 2026).

Table 4.7 Descriptive inclusion comparisons

Top-two measure	High confidence n=83	Lower confidence n=64	Barrier n=25	No named barrier n=122
Speed	67.5%	76.6%	64.0%	73.0%
Reliability	60.2%	56.3%	52.0%	59.8%
Convenience	65.1%	73.4%	68.0%	68.9%
Effectiveness	54.2%	57.8%	60.0%	54.9%
Satisfaction	55.4%	54.7%	60.0%	54.1%
Strong trust	34.9%	39.1%	28.0%	38.5%
High privacy concern	31.3%	42.2%	40.0%	35.2%
Ease	63.9%	71.9%	64.0%	68.0%

Source: Author analysis. Comparisons are descriptive and uncontrolled; high privacy concern is the top two concern categories.

4.5 Effects on the Online Shopping Experience

Expectations and satisfaction were similarly modest: 54.4% reported that support met UK-online-retailer expectations and 55.1% were satisfied, while their association was positive but weak ($\rho=.207$). The result is consistent with partial expectation confirmation, but prior expectations and journey stage were not measured as explanatory variables. AI cannot therefore be identified as the cause of the reported experience (Nguyen, Chiu and Le, 2021; Bell, Bryman and Harley, 2022).

Human recovery was central when automation failed. Human access was highly important to 83.0%, yet reported outcomes included successful and unsuccessful transfer, no available adviser and uncertainty about access. Task severity and adviser availability may partly explain those outcomes, while the questionnaire did not test whether context was preserved. The evidence supports visible recovery routes but not a universal handover-success estimate (Ozuem et al., 2024; Rese and Witthohn, 2025).

Successful resolution was the leading single influence on overall evaluation, ahead of speed and reliability. Trust, privacy and ease were selected less often, but a forced single choice does not make them irrelevant. The critical judgement is that shopping experience depends on task-relevant, dependable and recoverable service, not merely immediate response (Chen et al., 2022; Ozuem et al., 2024; Rese and Witthohn, 2025).

Table 4.8 Human access, influence and improvement priorities

Type	Category	Count	Percent
Q17 outcome	Human assistance not needed	47	32.0%
Q17 outcome	Transfer successful	34	23.1%
Q17 outcome	Transfer unsuccessful	28	19.0%
Q17 outcome	No human adviser available	18	12.2%
Q17 outcome	Unsure how to reach a human adviser	11	7.5%
Q18 influence	Successful problem resolution	38	25.9%
Q18 influence	Response speed	31	21.1%
Q18 influence	Reliability of information or action	28	19.0%
Q19 priority	Easier access to a human adviser	38	25.9%
Q19 priority	Better problem resolution	34	23.1%
Q19 priority	More accurate and reliable information	29	19.7%

Source: Author analysis of Q17–Q19 (n=147). Selected leading categories are shown; full distributions are available in the analysis workbook.

Human-access outcomes

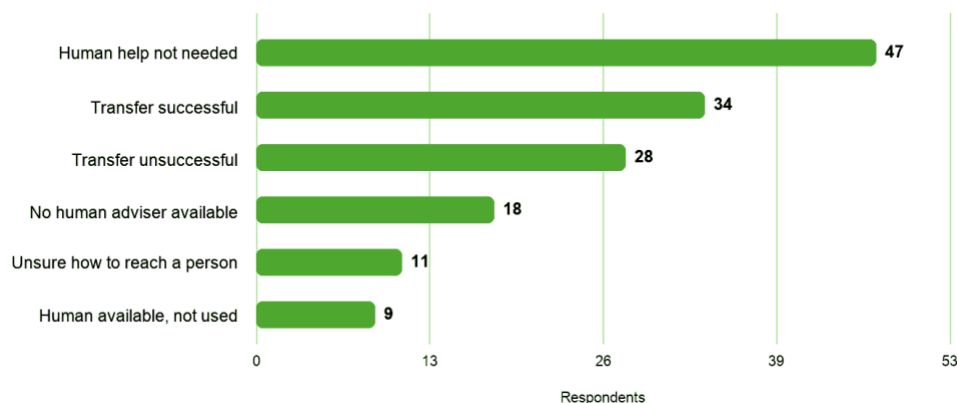


Figure 4.5 Human-access outcomes

Source: Author analysis of Q17 (n=147). Categories describe different circumstances and are not combined into one transfer-success denominator.

UK contextual sources establish the economic importance of internet retail and continuing concerns around delivery, redress and uneven digital participation, but they do not validate the survey's customer-level relationships (ONS, 2026a; Ofcom, 2025; Citizens Advice, 2025; Lloyds Banking Group, 2025; Ofcom, 2026). Their role is contextual triangulation: AI support should be governed as one accountable component of the shopping journey, with inclusive alternatives and human recovery when automated service is insufficient.

4.6 Practical Recommendations for Responsible and Effective AI Use

Respondents prioritized easier human access, better problem resolution and more accurate or reliable information ahead of faster responses. The ordering is coherent with the weaker effectiveness, satisfaction and trust findings and the mixed handover outcomes.

The evidence therefore favors resolution, reliability, inclusion and recovery before further speed optimization. This is a bounded managerial judgement: Q19 reports preferences, while service-quality and recovery literature explains why the priorities matter; neither establishes a universal intervention effect (Chen et al., 2022; Mubarak and Suomi, 2022; Ozuem et al., 2024).

Responsible use requires bounded automation, calibrated trust and proportionate data practice. Retailers should explain capability and data use, avoid implying unverified competence, test accessibility beyond average users and preserve accountable recovery after failure. Context-preserving transfer is a survey-derived testing recommendation, not a finding established by the retained recovery studies (Liao and Sundar, 2022; Ding and Najaf, 2024; Hasal et al., 2021; Rese and Witthohn, 2025).

4.7 Integrated Discussion

4.7.1 Pre-specified Associations

No selected coefficient exceeded .225 in absolute value, and six of eleven bootstrap intervals crossed zero. Speed-convenience and satisfaction-expectations showed weak positive evidence; effectiveness-satisfaction remained uncertain. Retaining negligible, negative and imprecise results prevents selective confirmation of the Hybrid Model and limits any claim that one indicator represents overall customer value (Bell, Bryman and Harley, 2022).

Table 4.9 Pre-specified Spearman associations with uncertainty intervals

Rank	Pair	rho	95% bootstrap CI	Magnitude
1	Response speed × convenience	0.225	0.064 to 0.374	Weak
2	Satisfaction × met expectations	0.207	0.046 to 0.356	Weak
3	Service quality × privacy concern	-0.176	-0.331 to -0.011	Weak
4	Digital confidence × convenience	-0.169	-0.322 to -0.008	Weak
5	Ease of use × human-adviser importance	0.167	0.005 to 0.326	Weak
6	Effectiveness × satisfaction	0.131	-0.035 to 0.303	Weak
7	Met expectations × service quality	-0.108	-0.258 to 0.048	Weak
8	Service quality × satisfaction	-0.102	-0.264 to 0.063	Weak
9	Digital confidence × ease of use	-0.089	-0.243 to 0.077	Negligible
10	Trust × privacy concern	0.026	-0.135 to 0.177	Negligible
11	Trust × satisfaction	-0.022	-0.178 to 0.134	Negligible

Source: Author analysis (n=147). Spearman coefficients use ordered survey items; intervals are 95% percentile bootstrap intervals from 5,000 resamples using the recorded deterministic seed sequence. They are descriptive and non-causal.

4.7.2 Triangulation of Primary and Secondary Evidence

The responsiveness expectation received partial support. Speed was rated more favorably than effectiveness, service quality and satisfaction, and its largest association was with convenience rather than a completed outcome. This converges with Chen et al. (2022), who separates responsiveness from reliability and problem solving. Different task difficulty and prior response-time expectations could also generate the gap, while single items cannot identify the mechanism. The finding is managerially important because a fast first response may improve perceived effort even when authority, information quality or recovery is insufficient. Retailers should therefore monitor speed alongside accuracy, resolution, repeat contact and the consequence of unresolved cases.

The trust-risk expectation remained unresolved. Strong trust was limited and privacy concern material, yet trust was negligibly related to both satisfaction and privacy concern. Liao and Sundar (2022) and Hasal et al. (2021) justify treating reliance and data risk as distinct, but the observed pattern may also reflect heterogeneous task sensitivity or self-selection by customers willing to use AI support. A low-risk tracking query and a refund dispute can carry different requirements even when both receive the same trust scale. Satisfaction cannot serve as evidence of justified trust; disclosure, data use, system authority and audited accuracy require separate assessment.

Expectation confirmation showed weak convergence rather than a complete explanatory pathway. Satisfaction and met expectations had similar positive rates and $\rho=.207$, consistent with Nguyen, Chiu and Le (2021), while high importance attached to human access supports recovery literature (Ozuem et al., 2024; Rese and Witthohn, 2025). Unmeasured prior expectations and journey-stage severity offer rival explanations, and a respondent may be satisfied with a modest outcome while still expecting better service from the retailer. Retailers should measure prior standards, post use satisfaction, verified resolution and recovery outcomes separately before attributing improvement to AI.

The enabling-conditions expectation was challenged by mixed inclusion evidence. Higher confidence did not correspond to consistently better convenience or ease, and the barrier subgroup combined weaker speed, reliability and trust with slightly higher effectiveness and satisfaction. This does not overturn the broader mechanisms identified by Mubarak and Suomi (2022) and Jha et al. (2026): small heterogeneous groups and online self-selection may obscure disadvantage. Aggregating different barriers may also conceal opposite effects, while higher-confidence customers may apply more demanding standards. Confidence and average ease are therefore insufficient proxies; accessibility must be tested by task, device, ability, assistance need and recovery route.

Across the four expectations, the Hybrid Model is most defensible as a diagnostic checklist rather than a predictive chain. Convergence supports separate task level measures and visible recovery; contradiction prevents satisfaction, confidence or speed from standing for responsible customer value. The model's value lies in exposing compromises and missing evidence: favorable usability cannot cancel a consequential error, and an average positive rating cannot demonstrate inclusion. This interpretation goes beyond asking whether AI was positively rated by asking which evidence would justify reliance for a specific service function and which adverse outcome would require human authority. Because the evidence is cross-sectional and self-reported, retailers should test the framework against verified accuracy, resolution, repeat contact, accessibility and redress data, segmented by task authority and customer need, before making deployment decisions.

Table 4.10 Integrated evidence and managerial implications

Theoretical expectation	Survey finding	Literature comparison	Critical judgement	Retail implication
Responsiveness is not dependable service	Speed 71.4%; satisfaction 55.1%; speed-convenience $\rho=.225$	Service quality separates responsiveness, reliability and problem solving	Partial support; task difficulty and prior expectations remain alternatives	Monitor time, accuracy, resolution and repeat contact separately
Trust-risk is distinct from satisfaction	Strong trust 36.7%; privacy concern 36.1%; trust relationships negligible	Calibrated trust and privacy are separate conditions of reliance	Unresolved; task sensitivity, single items and self-selection may attenuate links	Assess disclosure, data use, accuracy and satisfaction separately

Theoretical expectation	Survey finding	Literature comparison	Critical judgement	Retail implication
Expectation confirmation only partly explains satisfaction	Met expectations 54.4%; satisfaction 55.1%; $\rho=.207$	Expectation confirmation predicts comparison with prior standards	Weak convergence; prior expectations and journey severity were unmeasured	Measure prior standard, post-use judgement and recovery outcome
Confidence alone does not establish inclusion	High-confidence ratings were not consistently better; barrier results mixed	Inclusion research identifies access, skills, ability and recovery barriers	Challenged but uncertain because of small groups and online self-selection	Test by access need, device and task; retain assisted alternatives
Failure requires accountable human recovery	Human access highly important to 83.0%; handover outcomes mixed	Recovery research links post-failure response with customer outcomes	Convergent need, but transfer quality and task severity were not controlled	Provide visible transfer with retained context and sufficient authority

Source: Author triangulation of the survey and retained 40-source evidence corpus (Chen et al., 2022; Ozuem et al., 2024; Rese and Witthohn, 2025).

4.8 Chapter Summary

The observations address the revised objectives without claiming strong AI effects. Chatbots were the leading identified tool, while tracking and returns were leading tasks. Speed and convenience were the strongest performance indicators, but effectiveness, service quality and satisfaction were more mixed. Potential advantages coexisted with limited strong trust, material privacy concern and uneven participation evidence. The described online-shopping experience depended particularly on expectations, resolution and human recovery. All selected associations were weak or negligible. The combined literature supports responsible, task-bounded, inclusive and recoverable AI service (Chen et al., 2022; Liao and Sundar, 2022; Ozuem et al., 2024).

Chapter 5: Conclusions, Limitations and Recommendations

5.1 Introduction

This chapter answers the five revised research questions, evaluates achievement of the aim and converts the bounded evidence into practical guidance. The conclusions apply to the 147 person online convenience sample and retained evidence corpus. They do not estimate views among all UK shoppers, verify every retailer's technology use or establish causal AI effects on the online-shopping experience.

5.2 Answers to the Research Questions

5.2.1 RQ1: Main AI tools in customer support

Chatbots were the leading reported tool, followed by tracking/delivery support and AI help-center or search tools. The same tool category served different tasks and journey stages, so the evidence identifies customer-reported patterns rather than retailer market share or verified system classifications (Mariani, Hashemi and Wirtz, 2023; Nicolescu and Tudorache, 2022).

5.2.2 RQ2: Service performance and customer satisfaction

Speed and convenience were evaluated more favorably than effectiveness, service quality and satisfaction. Weak associations and overlapping exploratory group intervals show that no tool can be ranked as generally superior. RQ2 is therefore answered through separate, task-level performance judgements rather than a general AI-benefit claim (Chen et al., 2022; Ozuem et al., 2024).

5.2.3 RQ3: Benefits, limitations and digital exclusion

The combined evidence shows conditional benefits in relevance, convenience, ease and independent completion, alongside limited strong trust, material privacy concern and mixed inclusion evidence. Trust, privacy, accessibility and recovery cannot be inferred from speed or satisfaction and require direct evaluation (Liao and Sundar, 2022; Mubarak and Suomi, 2022).

5.2.4 RQ4: Effects on the online-shopping experience

Reported shopping experience depended on task relevance, expectations, dependable resolution and recoverable failure. Human access was strongly valued, but handover routes were uneven. These findings describe associations and evaluations within the achieved sample, not a causal effect of AI (Nguyen, Chiu and Le, 2021; Ozuem et al., 2024; Rese and Witthohn, 2025).

5.2.5 RQ5: Responsible and effective retailer action

UK online retailers should bound automation by task and authority, measure speed separately from accuracy and resolution, explain data use, test inclusion and provide visible human recovery. These actions combine respondent priorities with the retained service-quality, trust, inclusion and recovery evidence; they require retailer-specific baselines rather than universal targets (Liao and Sundar, 2022; Ozuem et al., 2024).

5.3 Achievement of the Aim and Contribution

The aim was achieved through five aligned research questions and three bounded contributions. Empirically, the study documents how 147 eligible respondents evaluated recent AI support across tools, tasks, performance, satisfaction, trust, privacy, inclusion and recovery. This contribution is descriptive and sample-specific rather than a national estimate.

Theoretically, the study conditionally evaluates the Hybrid Model. Responsiveness was distinct from dependable service, trust-risk relationships were unresolved, expectation confirmation was weak and confidence did not capture inclusion consistently. Preserving contradiction and uncertainty supports a diagnostic framework rather than a validated causal model (Bell, Bryman and Harley, 2022).

Practically, the study converts survey-literature convergence into task-level monitoring and human-recovery guidance. Retailers can test speed, accuracy, resolution, satisfaction, repeat contact, accessibility and redress separately, with accountable owners and local baselines. The proposed indicators operationalize responsible testing; they are not universal performance effects (Ozuem et al., 2024; Rese and Witthohn, 2025).

5.4 Recommendations for UK Online Retailers

Recommendations are limited to actions supported by the survey, retained literature or both. Retailers should establish task-specific baselines and examine adverse outcomes and customer needs alongside averages. The indicators are governance measures, not universal targets (Chen et al., 2022; Liao and Sundar, 2022).

Table 5.1 Evidence-linked recommendations for UK online retailers

Priority	Action	Evidence	Accountable function	Measurable indicator
1	Bound automation by task and authority	Both	Customer operations + AI product owner	% tasks resolved correctly; repeat-contact rate; exceptions requiring override
2	Make human access visible and effective	Both	Customer-service operations	Transfer availability; successful handover; time to authorized resolution
3	Measure speed, accuracy and resolution separately	Survey	Service quality + analytics	Response time; audited answer accuracy; end-to-end resolution; repeat contact
4	Transfer context with the customer	Both	CRM / platform owner	% handovers with transcript, intent and prior actions; repetition complaints
5	Use proportionate data and clear explanation	Both	Data protection officer + product	Data fields used; explanation visibility; deletion/objection completion
6	Test inclusion beyond average ease	Both	Accessibility lead + UX research	Task completion by access need/device; assisted-use failures; alternative-channel success

Source: Author recommendations derived from the survey and retained secondary evidence. Indicators require retailer-specific baselines and governance (Chen et al., 2022; Liao and Sundar, 2022; Mubarak and Suomi, 2022).

5.5 Limitations

Convenience recruitment through Discord, Facebook and X leaves coverage and selection probabilities unknown. Openings, refusals and abandonments were not retained, so response rate cannot be calculated; no a priori power target or separate pilot report is available. Online recruitment may underrepresent those facing the greatest exclusion, and pooled comparisons cannot control retailer, system version, task severity or customer need (Bell, Bryman and Harley, 2022; Saunders, Lewis and Thornhill, 2023; Mubarak and Suomi, 2022).

Cross-sectional self-report cannot establish temporal order, cause or verified operational outcomes. Single items constrain construct validity; Wilson and bootstrap intervals expose but do not remove sampling uncertainty. The structured review records query logic and appraisal fields but lacks historical result and exclusion counts. These boundaries preclude national prevalence, market-share claims, causal regression and a validated composite experience score (Bell, Bryman and Harley, 2022).

5.6 Future Research

Future research should use probability-informed or stratified UK recruitment and deliberately include lower-confidence, older, disabled and connectivity constrained participants. Validated multi-item measures, longitudinal or experimental designs and service log linkage could compare perceived experience with verified accuracy, resolution, escalation, repeat contact and remedy time across task severity and customer need (Bell, Bryman and Harley, 2022; Mubarak and Suomi, 2022).

5.7 Final Conclusion

AI support was often fast, convenient and easy to use in the achieved sample, but the evidence does not establish a strong, uniform or causal benefit. Chatbots were the main identified tool; service quality, effectiveness and satisfaction were more restrained than speed; strong trust was limited; privacy concern remained material; and effective human recovery was strongly valued. The online-shopping experience depended on more than immediate response, especially task relevance, dependable resolution and access to accountable assistance. Responsible and effective UK online retail use therefore requires task-bounded automation, separate performance and satisfaction evidence, proportionate data practice, inclusive alternatives and effective escalation. Adoption is an input; customer value must be demonstrated (Chen et al., 2022; Liao and Sundar, 2022; Ozuem et al., 2024).

References

- Adam, M., Wessel, M. and Benlian, A. (2021) 'AI-based chatbots in customer service and their effects on user compliance', *Electronic Markets*, 31(2), pp. 427-445. doi: <https://doi.org/10.1007/s12525-020-00414-7>. (Accessed: 23 June 2026).
- Alagarsamy, S. and Mehroliya, S. (2023) 'Exploring chatbot trust: Antecedents and behavioural outcomes', *Heliyon*, 9(5), e16074. doi: <https://doi.org/10.1016/j.heliyon.2023.e16074>. (Accessed: 23 June 2026).
- Alharbi, N., Ud Din, F., Paul, D. and Sadgrove, E. (2025) 'Driving AI chatbot adoption: A systematic review of factors, barriers, and future research directions', *Journal of Open Innovation: Technology, Market, and Complexity*, 11(3), 100590. doi: <https://doi.org/10.1016/j.joitmc.2025.100590>. (Accessed: 23 June 2026).
- Auer, I., Schlögl, S. and Glowka, G. (2024) 'Chatbots in airport customer service: Exploring use cases and technology acceptance', *Future Internet*, 16(5), 175. doi: <https://doi.org/10.3390/fi16050175>. (Accessed: 23 June 2026).
- Bahaw, P., Forgenie, D., Sadiq, G. and Sookhai, S. (2025) 'Generative AI for business sustainability: Examining usability, usefulness, and triple bottom line impacts in small and medium enterprises', *Sustainable Futures*, 10, 100815. doi: <https://doi.org/10.1016/j.sft.2025.100815>. (Accessed: 23 June 2026).
- Bell, E., Bryman, A. and Harley, B. (2022) *Business Research Methods*. 6th edn. Oxford: Oxford University Press.
- Beyari, H. (2025) 'Assessing artificial intelligence's impact on e-customer loyalty in the Saudi Arabian market', *Frontiers in Artificial Intelligence*, 8, 1541678. doi: <https://doi.org/10.3389/frai.2025.1541678>. (Accessed: 23 June 2026).
- Chau, H.K.L., Ngo, T.T.A., Bui, C.T. and Tran, N.P.N. (2025) 'Human-AI interaction in e-commerce: The impact of AI-powered customer service on user experience and decision-making', *Computers in Human Behavior Reports*, 19, 100725. doi: <https://doi.org/10.1016/j.chbr.2025.100725>. (Accessed: 23 June 2026).
- Chen, Q., Gong, Y., Lu, Y. and Tang, J. (2022) 'Classifying and measuring the service quality of AI chatbot in frontline service', *Journal of Business Research*, 145, pp. 552-568. doi: <https://doi.org/10.1016/j.jbusres.2022.02.088>. (Accessed: 23 June 2026).
- Citizens Advice (2025) Return to sender: It's time Ofcom sent a stronger message to delivery companies. Available at: <https://wearecitizensadvice.org.uk/return-to-sender-its-time-ofcom-sent-a-stronger-message-to-delivery-companies-be5778269464> (Accessed: 23 June 2026).
- Data Protection Act 2018, c. 12. London: The Stationery Office. Available at: <https://www.legislation.gov.uk/ukpga/2018/12/contents> (Accessed: 27 June 2026).

- Ding, Y. and Najaf, M. (2024) 'Interactivity, humanness, and trust: A psychological approach to AI chatbot adoption in e-commerce', *BMC Psychology*, 12(1), 595. doi: <https://doi.org/10.1186/s40359-024-02083-z>.
(Accessed: 23 June 2026).
- Hasal, M., Nowaková, J., Saghair, K.A., Abdulla, H., Snášel, V. and Ogiela, L. (2021) 'Chatbots: Security, privacy, data protection, and social aspects', *Concurrency and Computation: Practice and Experience*, 33(19), e6426. doi: <https://doi.org/10.1002/cpe.6426>.
(Accessed: 23 June 2026).
- Kumari, P., Gupta, N. and Aeron, P. (2025) 'The rise of the phygital citizens: Generative AI experiences and customer citizenship behaviour', *Journal of Global Information Management*, 33(1), pp. 1-17. doi: <https://doi.org/10.4018/JGIM.373320>.
(Accessed: 23 June 2026).
- Liao, Q.V. and Sundar, S.S. (2022) 'Designing for responsible trust in AI systems: A communication perspective', 2022 ACM Conference on Fairness, Accountability, and Transparency, pp. 1257-1268. doi: <https://doi.org/10.1145/3531146.3533182>.
(Accessed: 23 June 2026).
- Lloyds Banking Group (2025) Consumer Digital Index 2025: The Generative AI Edition. Available at: <https://www.lloydsbankinggroup.com/assets/pdfs/media/consumer-digital-index/2025/2025-consumer-digital-index.pdf>
(Accessed: 27 June 2026).
- Lopes, J.M., Silva, L.F. and Massano-Cardoso, I. (2024) 'AI meets the shopper: Psychosocial factors in ease of use and their effect on e-commerce purchase intention', *Behavioral Sciences*, 14(7), 616. doi: <https://doi.org/10.3390/bs14070616>.
(Accessed: 23 June 2026).
- Lu, Y. and Zhang, J. (2025) 'Balancing identity diversity and product contexts: Understanding consumer trust in AI-enhanced chatbot services', *Journal of Retailing and Consumer Services*, 84, 104205. doi: <https://doi.org/10.1016/j.jretconser.2024.104205>.
(Accessed: 23 June 2026).
- Mariani, M.M., Hashemi, N. and Wirtz, J. (2023) 'Artificial intelligence empowered conversational agents: A systematic literature review and research agenda', *Journal of Business Research*, 161, 113838. doi: <https://doi.org/10.1016/j.jbusres.2023.113838>.
(Accessed: 23 June 2026).
- Jha, G., Wright, J., Singhal, A., Zhang, Y., Burton, J. and McColl-Kennedy, J.R. (2026) 'Addressing vulnerability in customer experience with AI-agents', *Journal of Service Management*, 37(3), pp. 418-450. doi: <https://doi.org/10.1108/JOSM-04-2025-0182>.
(Accessed: 23 June 2026).
- McKinsey & Company (2025) The state of AI in 2025: Agents, innovation, and transformation. McKinsey Global Publishing. Available at: <https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai>
(Accessed: 23 June 2026).

- Meng, H., Lu, X. and Xu, J. (2025) 'The impact of chatbot response strategies and emojis usage on customers' purchase intention: The mediating roles of psychological distance and performance expectancy', *Behavioral Sciences*, 15(2), 117. doi: <https://doi.org/10.3390/bs15020117>.
(Accessed: 23 June 2026).
- Mubarak, F. and Suomi, R. (2022) 'Elderly forgotten? Digital exclusion in the information age and the rising grey digital divide', *INQUIRY: The Journal of Health Care Organization, Provision, and Financing*, 59. doi: <https://doi.org/10.1177/00469580221096272>.
(Accessed: 23 June 2026).
- Nguyen, D.M., Chiu, Y.-T.H. and Le, H.D. (2021) 'Determinants of continuance intention towards banks' chatbot services in Vietnam: A necessity for sustainable development', *Sustainability*, 13(14), 7625. doi: <https://doi.org/10.3390/su13147625>.
(Accessed: 23 June 2026).
- Nguyen, T.H. and Le, X.C. (2025) 'Artificial intelligence-based chatbots: A motivation underlying sustainable development in banking: standpoint of customer experience and behavioral outcomes', *Cogent Business & Management*, 12(1), 2443570. doi: <https://doi.org/10.1080/23311975.2024.2443570>.
(Accessed: 23 June 2026).
- Nicolescu, L. and Tudorache, M.T. (2022) 'Human-computer interaction in customer service: The experience with AI chatbots – A systematic literature review', *Electronics*, 11(10), 1579. doi: <https://doi.org/10.3390/electronics11101579>.
(Accessed: 23 June 2026).
- Office for National Statistics (ONS) (2026a) Internet sales as a percentage of total retail sales (ratio) (%), Great Britain: January 2026 (series J4MC). Available at: <https://www.ons.gov.uk/businessindustryandtrade/retailindustry/timeseries/j4mc/drsi>
(Accessed: 22 June 2026).
- Office for National Statistics (ONS) (2026b) Retail sales, Great Britain: May 2026. Available at: <https://www.ons.gov.uk/businessindustryandtrade/retailindustry/bulletins/retailsales/may2026>
(Accessed: 23 June 2026).
- Office for National Statistics (ONS) (2026c) Internet retail sales by sector, Great Britain, up to December 2025. Newport: Office for National Statistics. Available at: <https://www.ons.gov.uk/businessindustryandtrade/retailindustry/adhocs/3254internetretailsalesbysectorgreatbritainuptodecember2025>
(Accessed: 13 July 2026).
- Ofcom (2025) Best and worst parcel firms for customer satisfaction revealed. Office of Communications. Available at: <https://www.ofcom.org.uk/post/deliveries-and-charges/best-and-worst-parcel-firms-for-customer-satisfaction-revealed-2025>
(Accessed: 23 June 2026).
- Ofcom (2026) Adults' Media Use and Attitudes Report 2026. Office of Communications. Available at: <https://www.ofcom.org.uk/siteassets/resources/documents/research-and-data/media-literacy-research/adults/adults-media-use-and-attitudes-2026/adults-media-use-and-attitudes-2026-report.pdf>
(Accessed: 23 June 2026).

- Ozuem, W., Ranfagni, S., Willis, M., Salvietti, G. and Howell, K. (2024) 'Exploring the relationship between chatbots, service failure recovery and customer loyalty: A frustration–aggression perspective', *Psychology & Marketing*, 41(10), pp. 2253-2273. doi: <https://doi.org/10.1002/mar.22051>.
(Accessed: 23 June 2026).
- Rese, A. and Witthohn, L. (2025) 'Recovering customer satisfaction after a chatbot service failure – The effect of gender', *Journal of Retailing and Consumer Services*, 84, 104257. doi: <https://doi.org/10.1016/j.jretconser.2025.104257>.
(Accessed: 23 June 2026).
- Saunders, M., Lewis, P. and Thornhill, A. (2023) *Research Methods for Business Students*. 9th edn. Harlow: Pearson.
- Sidlauskiene, J., Joye, Y. and Auruskeviciene, V. (2023) 'AI-based chatbots in conversational commerce and their effects on product and price perceptions', *Electronic Markets*, 33(1), 24. doi: <https://doi.org/10.1007/s12525-023-00633-8>.
(Accessed: 23 June 2026).
- Uzoka, A., Cadet, E. and Ojukwu, P.U. (2024) 'Leveraging AI-powered chatbots to enhance customer service efficiency and future opportunities in automated support', *Computer Science & IT Research Journal*, 5(10), pp. 2485-2510. doi: <https://doi.org/10.51594/csitrj.v5i10.1676>.
(Accessed: 23 June 2026).
- Wadud, I., Liang, Y.D. and Polkinghorne, M. (2025) 'Digital transformation in the UK retail sector', *Encyclopedia*, 5(3), 142. doi: <https://doi.org/10.3390/encyclopedia5030142>.
(Accessed: 23 June 2026).
- Xu, Y., Zhang, J. and Deng, G. (2022) 'Enhancing customer satisfaction with chatbots: The influence of communication styles and consumer attachment anxiety', *Frontiers in Psychology*, 13, 902782. doi: <https://doi.org/10.3389/fpsyg.2022.902782>.
(Accessed: 23 June 2026).
- Yun, J. and Park, J. (2022) 'The effects of chatbot service recovery with emotion words on customer satisfaction, repurchase intention, and positive word-of-mouth', *Frontiers in Psychology*, 13, 922503. doi: <https://doi.org/10.3389/fpsyg.2022.922503>.
(Accessed: 23 June 2026).
- Zaheer, A., Sadiq, A. and Safi, N. (2025) 'AI in service industries: Effects on customer satisfaction, mediated by service quality, and moderated by customer trust', *Journal of Social Sciences and Humanities*, 2(3), pp. 51-73. doi: <https://doi.org/10.62810/jssh.v2i3.82>.
(Accessed: 23 June 2026).

Glossary of Terms

AI customer support: Customer-service assistance delivered partly or wholly through artificial-intelligence systems, including chatbots, virtual assistants, automated tracking, help search and task-oriented agents.

Chatbot: A conversational software interface that responds to customer messages using predefined rules, machine learning or generative AI.

Convenience sample: A non-probability sample recruited from participants who were accessible and willing to take part; it does not provide known selection probabilities.

Cross-sectional survey: A survey that records participants' responses at one point in time and cannot by itself establish causal or temporal relationships.

Digital inclusion: The ability to access, understand and use digital services effectively, including suitable devices, connectivity, skills, confidence and accessible design.

Human escalation: Transfer from automated support to an authorized human adviser when the automated system cannot complete or safely resolve the task.

Privacy concern: A respondent's concern about how personal or conversational data are collected, used, retained, shared or protected.

Service quality: The customer's assessment of service performance, including responsiveness, reliability, competence, effectiveness and resolution.

Service recovery: Action taken after a service failure to correct the problem, restore the customer's position and rebuild satisfaction or confidence.

Spearman's rho: A rank-based coefficient used here to describe the direction and strength of association between selected ordinal survey variables.

Trust: The extent to which a customer is willing to rely on an AI support system, subject to its demonstrated capability, transparency and limits.

DPA: Data Protection Act 2018.

GDPR: General Data Protection Regulation.

ONS: Office for National Statistics.

RQ: Research Question.

TAM: Technology Acceptance Model.

UTAUT: Unified Theory of Acceptance and Use of Technology.

Appendices

Appendix A: Administered Questionnaire

The following question wording and response categories reproduce the approved questionnaire used for the 3-31 July 2026 online data collection. Respondents were instructed to answer with reference to their most recent AI online-shopping support experience. The substantive questions and response options are unchanged; the section headings below omit the questionnaire's former objective labels because Appendix B provides the revised analytical mapping.

Participant Information and Consent

Participation is voluntary and should take approximately 5–7 minutes. Please do not provide your name, contact details, account information, payment details or other identifying information. Responses will be analyzed in aggregate. You may stop at any time before submitting the questionnaire.

Eligibility Questions

S1. Do you consent to take part in this questionnaire?

If No, please end the questionnaire.

Response options: Yes / No

S2. Are you aged 18 or over and currently resident in the UK?

If No, please end the questionnaire.

Response options: Yes / No

S3. During the last 12 months, have you shopped online and used or encountered AI customer support?

Examples include chatbots, virtual assistants, automated tracking, automated returns or AI shopping assistants.

If No or Not sure, please end the questionnaire.

Response options: Yes / No / Not sure

Participant Profile

P1. What is your age group?

Response options: 18–24 / 25–34 / 35–44 / 45–54 / 55–64 / 65 or over / Prefer not to say

P2. How often do you shop online?

Response options: Several times a week / About once a week / Several times a month / A few times a year / Rarely

P3. Overall, how confident are you in using digital services when shopping online?

Response options: Not at all confident / Slightly confident / Moderately confident / Very confident / Extremely confident

P4. Do any of the following make it more difficult for you to participate in online shopping or use online customer support?

Select all that apply. If none apply, select “No barriers” only.

Response options: No barriers / Limited or unreliable internet or device access / Cost of internet access or digital devices / Accessibility or disability-related needs / Language or literacy difficulties / Limited digital skills or confidence / Security or privacy concerns / Needing help from another person / Other / Prefer not to say

Section A: AI Support Tools and Customer-Service Tasks

Q1. Which AI support tool did you use in your most recent online-shopping support experience?

Response options: Text-based chatbot / Voice or virtual assistant / AI-powered help center or search tool / Automated order-tracking or delivery support / Automated returns, exchange or refund system / AI product-recommendation or shopping assistant / AI-generated email or messaging support / Other AI support / Not sure

Q2. What was the main customer-service task you wanted the AI support to handle?

Response options: Finding product information / Comparing products or receiving recommendations / Account access or login support / Payment or checkout support / Checking an order or delivery / Changing or cancelling an order / Returning or exchanging a product / Requesting a refund / Making a complaint / Resolving another problem / Other

Q3. At which stage of the shopping journey did the interaction take place?

Response options: Before purchase / During purchase or payment / After purchase: tracking or delivery / After purchase: return, exchange or refund / Complaint or problem resolution

Q4. How relevant was the AI tool to the customer-service task you needed to complete?

Response options: Not at all relevant / Slightly relevant / Moderately relevant / Very relevant / Extremely relevant

Section B: Support Performance

Q5. How quickly did the AI support respond?

Response options: Very slowly / Slowly / Neither slowly nor quickly / Quickly / Very quickly

Q6. How reliable was the AI support in providing information or actions you could depend on?

Response options: Very unreliable / Unreliable / Neither unreliable nor reliable / Reliable / Very reliable

Q7. How convenient was it to use the AI support for this customer-service task?

Response options: Very inconvenient / Inconvenient / Neither inconvenient nor convenient / Convenient / Very convenient

Q8. How effectively did the AI support resolve or complete your customer-service task?

Response options: Very ineffectively / Ineffectively / Neither ineffectively nor effectively / Effectively / Very effectively

Q9. How would you rate the overall service quality of the AI support?

Response options: Very poor / Poor / Neither poor nor good / Good / Excellent

Section C: Satisfaction, Trust, Privacy and Digital Inclusion

Q10. Overall, how satisfied were you with the AI support?

Response options: Very dissatisfied / Dissatisfied / Neither dissatisfied nor satisfied / Satisfied / Very satisfied

Q11. How much did you trust the AI support to provide accurate and appropriate information or assistance?

Response options: Not at all / Slightly / Moderately / Very much / Completely

Q12. How concerned were you about how the AI support collected, used or stored personal data?

Response options: Not at all concerned / Slightly concerned / Moderately concerned / Very concerned / Extremely concerned

Q13. How easy was the AI support to access, understand and use?

Response options: Very difficult / Difficult / Neither difficult nor easy / Easy / Very easy

Q14. Were you able to complete the AI-supported interaction without help from another person?

Response options: Yes, completely / Yes, mostly / Only partly / No / Prefer not to say

Section D: UK E-commerce Context and Consumer Expectations

Q15. To what extent did the AI support meet your expectations of an online retailer serving UK customers?

Response options: Definitely did not / Probably did not / Not sure / Probably did / Definitely did

Q16. If AI support cannot resolve a problem, how important is access to a human adviser?

Response options: Not at all important / Slightly important / Moderately important / Very important / Extremely important

Q17. What happened regarding access to a human adviser during your most recent interaction?

Response options: Human assistance was not needed / A human adviser was available and the transfer was successful / A human adviser was available, but the transfer was unsuccessful / A human adviser was available, but I did not use that option / No human adviser was available / I was not sure how to reach a human adviser

Q18. Which factor had the greatest influence on your overall evaluation of the AI support?

Response options: Response speed / Reliability of the information or action / Convenience / Successful problem resolution / Ease of use or accessibility / Trust in the AI support / Privacy or personal-data concerns / Access to a human adviser / Other

Section E: Responsible Customer-Centered Use

Q19. Which improvement should UK online retailers priorities when using AI customer support?

Response options: More accurate and reliable information / Faster responses / Better problem resolution / Easier and more accessible interfaces / Clearer explanations of AI use / Stronger privacy and data protection / Easier access to a human adviser / Better transfer of information from AI to human advisers / Other

Appendix B: Objective and Question Mapping

Table B1 Question-to-objective mapping

Objective	Exact objective	Question coverage
1	To identify the main AI tools used by UK online retailers for customer support.	Q1 primary; Q2-Q4 provide task, stage and relevance context
2	To assess how AI tools influence service performance, including response speed, service quality, and customer satisfaction.	Q5-Q10: speed, reliability, convenience, effectiveness, quality and satisfaction
3	To explore the benefits and limitations of AI support, including trust, privacy, and digital exclusion.	P3-P4; Q4 and Q6-Q14: benefits, trust, privacy,

		ease, completion and barriers
4	To examine how AI customer support affects the online shopping experience.	Q3, Q10 and Q15-Q18: stage, satisfaction, expectations, resolution and human access
5	To provide practical recommendations for UK online retailers on responsible and effective AI use.	Q19 plus triangulation across all findings

Source: Author revised mapping; objective wording reproduced exactly as specified, with the approved questions and response categories unchanged.

Appendix C: Coding and Analysis Protocol

Table C1 Coding protocol

Component	Rule	Boundary
Raw data	Answers worksheet A1:AJ148	Preserved unchanged in the analysis workbook
Ordered items	Q4–Q13 and Q15–Q16 coded in questionnaire order	Higher Q12 means greater privacy concern
Q14	No=1; Only partly=2; Yes mostly=3; Yes completely=4	Prefer not to say unavailable; valid n=140
Barriers	Any named P4 barrier=1	No named barrier=0; not a complete exclusion measure
Primary statistics	Counts, percentages, distributions, medians	No composite scale, alpha, factor analysis or regression
Associations	Eleven pre-specified Spearman coefficients	Retain weak, null, negative and unexpected results
Groups	High/lower confidence; barrier/no named barrier	Descriptive only; no population inference

Source: Author analysis protocol reconciled with Survey-Analysis.xlsx.

Appendix D: Data-Quality Audit

Table D1 Data-quality audit

Check	Result	Status
Eligibility	147/147 Yes for consent, UK resident aged 18+, and recent AI-support exposure	Pass
Completeness	No missing analyzed values across A1:AJ148	Pass
Identifiers	147 unique Record_ID values	Pass
Duplicate patterns	No duplicate complete answer patterns	Pass
Categories	All values matched administered categories	Pass
P4 logic	No No-barriers selection combined with a named barrier	Pass
Q14	Seven Prefer not to say retained; excluded only from ordinal analysis	Pass
Formula scan	No REF, DIV/0, VALUE, NAME or N/A errors in final workbook	Pass

Source: Author reconciliation of the original raw workbook and final analysis workbook.

Appendix E: Secondary-Evidence Search and Appraisal

The review retained 40 academic, official, public and limited industry sources relevant to the five objectives. Searches during June and July 2026 used Google Scholar, UWTSD library resources, publisher pages and official UK websites. The thematic query structure combined AI chatbot or conversational agent terms with online retail/e-commerce and service quality, satisfaction, trust, privacy, accessibility, digital inclusion, recovery or human escalation. UK context searches added ONS, Ofcom, parcel delivery, digital participation and consumer redress terms.

Sources were appraised by method/design, sector/geography, evidential role and principal transfer limitation. Peer-reviewed studies supported mechanisms and construct interpretation; official sources supported UK operating context; industry evidence was restricted to current capability or organizational adoption. Evidence was not pooled across incompatible designs, and non-UK or non-retail findings were not treated as direct UK effect estimates.

Historical search-result and exclusion counts were not retained, so this remains a structured bounded review rather than a systematic review. The bibliographic audit corrected the research methods textbook to Bell, Bryman and Harley (2022), sixth edition; the corpus remains 40 sources.

Table E1 Secondary-evidence selection and appraisal

ID / source	Method or design	Sector / geography	Evidential role	Principal transfer limitation
1. Adam, M., Wessel, M. and Benlian, A. (2021)	Randomized online experiment	E-commerce/customer service; Germany/Denmark-linked authorship	Anthropomorphic verbal cues, social presence and compliance	Does not test disclosure, accuracy or UK retail outcomes
2. Alagarsamy, S. and Mehroliia, S. (2023)	Quantitative chatbot-trust study	Banking chatbot; non-UK	Antecedents and outcomes of chatbot trust	Sector and geography limit transfer to UK retail
3. Alharbi, N., Ud Din, F., Paul, D. and Sadgrove, E. (2025)	Systematic literature review	Multi-sector international literature	Adoption factors and barriers	Review breadth does not establish UK customer prevalence
4. Auer, I., Schögl, S. and Glowka, G. (2024)	Airport chatbot use-case and acceptance study	Airport customer service; Austria	Usefulness, ease and acceptance mechanisms	Airport acceptance is not completed online retail support
5. Bahaw, P., Forgenie, D., Sadiq, G. and Sookhai, S. (2025)	SME survey on generative-AI usability and sustainability	SMEs; Trinidad and Tobago context	Usability/usefulness of generative AI in SMEs	SME adoption/sustainability is not evidence of customer-support scale or hours
6. Bell, E., Bryman, A. and Harley, B. (2022)	Business research-methods textbook	General business research	Research design, sampling and inference guidance	General methods guidance; not substantive AI-support evidence
7. Beyari, H. (2025)	Cross-sectional Saudi e-commerce survey	Saudi Arabian e-commerce	Associations among AI perceptions and e-loyalty	Cross-sectional, Saudi sample; no causal UK inference
8. Chau, H.K.L., Ngo, T.T.A., Bui, C.T. and Tran, N.P.N. (2025)	Quantitative e-commerce human-AI interaction study	E-commerce; non-UK sample	Human-AI interaction, user experience and decision-making	Non-UK and study-specific; cannot establish population effects
9. Chen, Q., Gong, Y., Lu, Y. and Tang, J. (2022)	Multi-study scale-development and validation research	Frontline AI-chatbot service; multi-study	Separate dimensions of AI-chatbot service quality	Construct validation does not prove retailer-level causal outcomes
10. Citizens Advice (2025)	UK consumer-policy commentary using parcel-delivery evidence	UK parcel delivery and redress	UK parcel-resolution and redress concerns	Public commentary is not peer-reviewed causal evidence
11. Data Protection Act (2018)	UK primary legislation	United Kingdom	Data Protection Act 2018 obligations	Does not by itself cite UK GDPR or show customer understanding
12. Ding, Y. and Najaf, M. (2024)	Quantitative e-commerce chatbot-adoption study	E-commerce; non-UK	Interactivity, humanness and trust in e-commerce	Non-UK cross-sectional/behavioural context
13. Hasal, M., Nowaková, J., Saghair, K.A., Abdulla, H., Snášel, V. and Ogiela, L. (2021)	Narrative review of chatbot security, privacy and social issues	Cross-sector chatbot governance	Data collection, security, privacy and social risks	Review-level governance evidence; not legal advice or outcome measurement
14. Kumari, P., Gupta, N. and Aeron, P. (2025)	Cross-sectional survey of 282 Prolific participants	Global online Prolific sample; phygital/generative AI	Generative-AI experience, trust and citizenship behaviour	One online global sample; not UK retail prevalence
15. Liao, Q.V. and Sundar, S.S. (2022)	Conceptual communication framework	Cross-sector responsible AI	Calibrated/responsible trust and communication	Conceptual, not an empirical effect estimate
16. Lloyds Banking Group (2025)	Mixed-source UK digital-inclusion report	UK digital participation; financially/digitally engaged and disengaged groups	Uneven UK access, confidence and digital participation	Digital/financial context, not online retail support outcomes
17. Lopes, J.M., Silva, L.F. and Massano-Cardoso, I. (2024)	Quantitative e-commerce purchase-intention study	E-commerce; non-UK	Ease of use and purchase-intention mechanisms	Purchase intention is not support resolution
18. Lu, Y. and Zhang, J. (2025)	Seven experiments on chatbot identity and product context	Retail chatbot experiments; South Korea-linked context	Identity cues, product context, trust and preference	Identity/product-context experiments do not cover inclusion or human handover

ID / source	Method or design	Sector / geography	Evidential role	Principal transfer limitation
19. Mariani, M.M., Hashemi, N. and Wirtz, J. (2023)	Systematic literature review	Conversational agents across sectors	Heterogeneity and research themes in conversational agents	Review categories do not prove the dissertation's full retail tool list
20. Jha, G., Wright, J., Singhal, A., Zhang, Y., Burton, J. and McColl-Kennedy, J.R. (2026)	Conceptual theory-adaptation paper	Personal AI agents and consumer vulnerability; conceptual, multi-country authorship	Consumer vulnerability, agency and personal AI-agent design	Conceptual personal-agent framework, not observed UK retail outcomes
21. McKinsey & Company (2025)	Global organizational survey (1,993 respondents)	Global organizations; not a customer-level retail study	Current organizational AI/agent adoption and customer-service automation	Organization-level and self-reported; not customer-level or retail-specific
22. Meng, H., Lu, X. and Xu, J. (2025)	Scenario/experimental study of response strategy and emojis	Online customer interaction; non-UK	Response strategy, psychological distance and purchase intention	Purchase intention and interface cues are not resolution evidence
23. Mubarak, F. and Suomi, R. (2022)	Review/conceptual paper on older-adult digital exclusion	Digital exclusion/health-information context; not retail-specific	Digital exclusion and the grey digital divide	Older-adult/digital-exclusion focus; not direct retail outcome evidence
24. Nguyen, D.M., Chiu, Y.-T.H. and Le, H.D. (2021)	Cross-sectional Vietnamese banking survey	Vietnamese banking	Expectation/continuance mechanisms in banking chatbots	Vietnamese banking and continuance, not UK retail
25. Nguyen, T.H. and Le, X.C. (2025)	Cross-sectional banking customer study	Banking; non-UK	Customer experience and behavioural outcomes in banking	Banking and cross-sectional; cannot prove causation
26. Nicolescu, L. and Tudorache, M.T. (2022)	Systematic literature review	Customer-service chatbots across sectors	Cross-study chatbot experience themes	Review evidence is heterogeneous and not UK-specific
27. Office for National Statistics (ONS) (2026a)	Official GB retail time series	Great Britain retail	Economic materiality of internet retail	Economic activity only; no AI support outcomes
28. Office for National Statistics (ONS) (2026b)	Official GB retail statistical bulletin	Great Britain retail	Current GB retail-sales context	Retail sales only; no AI support outcomes
29. Office for National Statistics (ONS) (2026c)	Official GB internet-retail sector dataset	Great Britain retail by sector	Sector distribution of GB internet retail	Sector sales only; no AI support outcomes
30. Ofcom (2025)	Official UK parcel-customer survey/news release	UK parcel delivery	Parcel-firm satisfaction and complaint context	Parcel evidence, not AI evidence
31. Ofcom (2026)	Official UK media-use survey report	UK adult media use and digital participation	UK digital participation and confidence	Participation context, not e-commerce service performance
32. Ozuem, W., Ranfagni, S., Willis, M., Salvietti, G. and Howell, K. (2024)	Empirical chatbot service-failure/recovery study	Chatbot service failure and loyalty; non-UK	Chatbot failure, recovery and loyalty	Recovery study does not necessarily test human escalation or context transfer
33. Rese, A. and Witthohn, L. (2025)	Pre-study (n=30) plus scenario experiment (n=300)	Chatbot service recovery; Germany	Recovery effects on warmth, competence, humanness and satisfaction	Recovery/gender experiment does not test direct human-vs-AI trust or digital inclusion
34. Saunders, M., Lewis, P. and Thornhill, A. (2023)	Business research-methods textbook	General business research	Business-research design and inference limits	General methods guidance, not evidence for substantive claims
35. Sidlauskienė, J., Joye, Y. and Aurskuskevičienė, V. (2023)	Experimental conversational-commerce study	Conversational commerce; non-UK	Product/price perceptions in conversational commerce	Product/price perceptions, not support completion
36. Uzoka, A., Cadet, E. and Ojukwu, P.U. (2024)	Narrative review of AI-chatbot service efficiency	Cross-sector review; not UK-retail primary evidence	Efficiency opportunities and 24/7 support claims at review level	Review quality and scope require cautious use; no causal UK estimate
37. Wadud, I., Liang, Y.D. and Polkinghorne, M. (2025)	Narrative review of UK retail digital transformation	UK retail; review article	Digital transformation themes in UK retail	Review-level evidence; limited direct customer measurement

ID / source	Method or design	Sector / geography	Evidential role	Principal transfer limitation
38. Xu, Y., Zhang, J. and Deng, G. (2022)	Quantitative study of chatbot style and attachment anxiety	Chatbot communication; non-UK	Communication style, attachment anxiety and satisfaction	Non-UK and specific constructs
39. Yun, J. and Park, J. (2022)	Experimental chatbot service-recovery study	Chatbot recovery; non-UK	Recovery language, satisfaction, repurchase and word of mouth	Scenario/experimental recovery context; transfer cautiously
40. Zaheer, A., Sadiq, A. and Safi, N. (2025)	Cross-sectional survey across service industries	Afghanistan-linked service sectors, especially healthcare/education	Surveyed associations among AI, service quality, trust and satisfaction	Broad service sectors and emerging journal; not UK retail

Source: Author critical appraisal of the retained 40-source evidence corpus.

Appendix F: Ethics Statement

The final ethics application records supervisor approval/sign-off by Dr Lakshani Mendis on 25 April 2026, before survey recruitment began. The protocol defined a low-risk online questionnaire for UK adults aged 18 or over with relevant online-shopping/customer-support experience. Participation was voluntary; an introductory page explained the purpose, procedures, expected time, anonymity and data use, and affirmative consent was required. Participants could omit questions or stop before submission, and a debrief offered project information and a findings summary on request.

The protocol anticipated possible inconvenience or mild discomfort when considering satisfaction, trust, privacy or negative support experiences, plus a limited online data-protection risk. Mitigation included avoiding highly sensitive questions, collecting no special-category data, using anonymized IDs, aggregate reporting, restricted password-protected University Office 365 storage, and deletion after project completion. No interviews, focus groups, observation, deception or audio/video recording were approved or used. The form contains no committee name or approval reference; the approval is therefore reported as supervisor sign-off.

Appendix G: Research Timeline

Table G1 Research timeline

Activity	June	July	August
Topic and literature development	X	X	
Questionnaire and primary-data preparation		X	
Data collection and validation		X	
Analysis and chart production		X	
Writing, integration and revision		X	X
Final checks and submission preparation			X

Source: Author research record and schedule.

Appendix H: AI-Use and Academic-Integrity Statement

Generative AI was used only for limited proofreading support, including identifying grammatical and syntactical issues in text written by the researcher. It was not used to generate research content, analyze data, produce charts, process spreadsheets, interpret findings, fabricate references, or make academic decisions. The dissertation, analysis and conclusions remain the researcher's own work and responsibility.