Available online at www.sciencedirect.com

ScienceDirect

Journal of Retailing 99 (2023) 440–459

www.elsevier.com/locate/jretai

Testing the performance of online recommendation agents: A meta-analysis

Markus Blut[a,*], Arezou Ghiassaleh[a], Cheng Wang[b]

[a] *Durham University Business School, Durham University, Mill Hill Lane, Durham DH1 3LB, UK*

[b] *International Business School, Zhejiang University, 718 East Haizhou Road, Haining, Zhejiang, 314400, China*

Available online 19 August 2023

Abstract

Many retailers (e.g., Amazon, Walmart) use various types of online recommendation agents (RAs) on their websites to suggest goods and services to consumers. These RAs screen millions of options to ease consumers' information search and evaluation. To determine which RA types best support consumers' efforts, the present research reports a meta-analysis of perceived recommendation quality research, a key performance metric that gauges RAs from consumers' perspectives. To test the framework derived from this meta-analysis, the authors rely on data gathered from 32,172 consumers, reported in 122 samples. The results affirm that some RAs perform better than others in leveraging the effects of perceived recommendation quality on consumers' decision-making satisfaction, RA satisfaction, and intention to use the RA in the future. The best performing RAs feature specific algorithms (i.e., collaborative filtering, interactive RAs, and self-serving recommendations), recommendation presentations (i.e., solicited recommendation), and data sources (i.e., location-based and social network–based RAs). Moreover, the results suggest that some RAs perform better than others in leveraging the effects of decision-making and RA satisfaction on future use intentions. These insights advance RA theory and provide guidance for managers, with regard to choosing the optimal RA.

Keywords: Meta-analysis; Online recommendation agent; Perceived recommendation quality; E-commerce; Technology acceptance; Consumer decision making.

Retailers aiming to expand their online business must manage and present product information effectively to their consumers (Marchand and Marx 2020). Consumers have increasingly easy access to various offerings, but that means that they must exert considerable cognitive effort to compare and evaluate the myriad choices (Sethuraman, Gázquez-Abad and Martínez-López 2022). For example, Amazon, together with its 85,000 Marketplace sellers, offers more than 350 million goods (Retailtouchpoints 2021). To ease information search and evaluation efforts, online retailers often use electronic recommendation agents (RAs), that is, "software agents that elicit the interests or preferences of individual consumers either explicitly or implicitly, and make recommendations accordingly" (Xiao and Benbasat 2007, pp. 137–38). Using the consumer's own stated preferences and past purchase behavior, or the purchase behavior of consumers with similar profiles, RAs recommend not only goods but also services (e.g., robo-advisors in banking, personalization of web searches). Managers can choose from several types of RA, though few guidelines exist to help them. A notable exception is McKinsey's (2020) report on the potential of novel RAs for firms, given recent technology advancements in artificial intelligence (AI).

To help managers select an RA, we investigate which RA types work best in supporting consumers, by turning to prior studies that identify variations in RAs' performance and effectiveness. Although all RAs use a range of algorithms and technologies to help consumers deal with information overload (Zhang, Agarwal and Lucas Jr 2011), not every RA is capable of supporting consumers (Marchand and Marx 2020). Extant studies tend to specify two RA types: collaborative or content-based filtering, depending on whether the RA relies on decisions made by similar consumers or the con-

* Corresponding author.
E-mail address: markus.blut@durham.ac.uk (M. Blut).

https://doi.org/10.1016/j.jretai.2023.08.001

sumer's own stated preferences and past purchase behavior (Ricci, Rokach and Shapira 2022). Yet other recommendation techniques are possible. For example, some RAs use data from social media (e.g., Facebook) or consumers' social networks. Others use location data (GPS) from mobile devices to improve predictions. Table 1 distinguishes three groups of characteristics that can define RA types: algorithm, recommendation presentation, and data source (Tsekouras, Li and Benbasat 2022). Each type leverages unique mechanisms, which implies they likely differ in how they shape consumers' decision-making satisfaction, satisfaction with the RA, and intentions to use the RA in the future. While most studies have examined single recommendation techniques, we seek to provide a comparative performance assessment.

Specifically, with a meta-analysis, we develop and test a framework of the performance of different RA types. We build this framework on the basis of research on perceived recommendation quality, a key performance metric for RAs that reflects the consumer's perspective (Tsekouras, Li and Benbasat 2022). *Perceived recommendation quality* is defined as the degree to which consumers perceive that the RA can assist them in making decisions (Nilashi et al. 2016). It is a broad construct that includes consumers' perceptions of an RA's recommendations as accurate, novel, and diverse (Nilashi et al. 2016). The framework considers perceived recommendation quality a key driver of consumers' *decision-making satisfaction, RA satisfaction*, and *future use intention*. Drawing on information diagnosticity theory, we explore how different RA types leverage the impact of perceived recommendation quality on these outcomes (Khare, Labrecque and Asare 2011).[1] The framework also includes potential mediating effects of consumers' decision-making and RA satisfaction on future use intentions. Thus, we explore which RA types best leverage the effects of these mediators (Seiders et al. 2005). Using our meta-analytic database, which includes data from 32,172 consumers in 122 samples, our study makes two major contributions to the literature.

First, we offer a performance assessment of RA types. The results suggest that perceived recommendation quality relates positively to consumers' decision-making satisfaction, RA satisfaction, and future use intentions. Furthermore, some RAs perform better than others in leveraging the effects of perceived recommendation quality on these outcomes. We find that the diagnosticity of perceived recommendation quality for consumers depends on the RA's underlying algorithm, recommendation presentation, and data sources (Khare, Labrecque and Asare 2011). Regarding the *RA algorithm*, we find that collaborative-filtering RAs, which generate recommendations on the basis of the interests expressed by similar consumers, enhance the effect of perceived recommendation quality. Self-serving recommendation RAs, which generate recommendations that often benefit firms/retailers more than consumers (e.g., "top picks"), also enhance some effects. Interactive RAs, which use novel AI technology to provide "in-the-moment" recommendations, perform well too. For *RA presentation*, we find only one high-performing RA type: those that rely on solicited recommendations, provided only after consumers' direct requests. Regarding *RA data*, we find that RAs that use consumers' location and social media data enhance some effects of perceived recommendation quality. These findings help clarify variations in RA performance.

Second, we explore how RA types leverage the effects of the two mediators in our framework, decision-making satisfaction and RA satisfaction, on consumers' intention to use the RA in the future. Both mediators relate positively to this outcome. Some RAs better enhance these effects than others, depending again on the RA's algorithm (e.g., collaborative filtering), recommendation presentation (e.g., solicited recommendation and comprehensiveness of results), and data source (e.g., personalized recommendation). Our results provide important insights into the role of RA types in translating consumers' satisfaction into future use intentions; they not only advance RA theory and guide scholars in studying promising RAs, but they also give managers insights into how to select an RA.

Conceptual background and literature review

Perceived recommendation quality

An RA's role is to suggest products of interest to consumers. Useful recommendations help consumers make better decisions and find suitable products, which can enhance their shopping experience. Perceived recommendation quality is a critical determinant of RA success (Tsekouras, Li and Benbasat 2022); Nilashi et al. (2016) define "recommendation quality" as the degree to which consumers perceive that the RA can assist them in making decisions, reflecting their perceptions of recommendations as accurate, novel, and diverse. Yet the operationalization of this construct has evolved. Initially, it was discussed as providing recommendations that fit a consumer's preferences (i.e., perceived recommendation accuracy; Herlocker et al. 2004). Being able to predict consumer preferences accurately has been the main driver of the RA field, and studies affirm its impact on consumers' intention to use the RA in the future (Tsekouras, Li and Benbasat 2022).

More recent research also acknowledges that perceived recommendation quality needs to go beyond accuracy and adopts a broader perspective. For example, McNee, Riedl and Konstan (2006, p. 1101) explain that "being accurate is not enough" and call for research on "recommenders from a user-centric perspective to make them not only accurate and helpful, but also a pleasure to use." That is, novelty and diversity also are important to the recommendation quality construct, as revealed by more comprehensive measures developed by researchers (Nilashi et al. 2016). Perceived recommendation

[1] Because most empirical studies examine single recommendation techniques, we cannot calculate effect sizes to assess the direct effects of RA types in the meta-analysis. For further information, see Borenstein et al. (2009, p. 3)), who explain, "The effect size, a value which reflects the magnitude of the treatment effect or (more generally) the strength of a relationship between two variables, is the unit of currency in a meta-analysis."

Table 1
Characteristics of different RA types and performance differences.

RA characteristics	Mechanism	Advantages	Disadvantages	Effect on decision-making satisfaction	Effect on RA satisfaction	Effect on future use intention	Sources
RA Algorithm							
Collaborative filtering	Generates recommendations based on interests of similar consumers	Credible recommendations and broad exposure to many products	Difficult to provide recommendations when users or products are new	More accurate recommendations enhance effect of perceived recommendation quality (RQ) on decision-making satisfaction	Higher confidence in RA recommendations enhances RQ effect on consumer satisfaction with RA	Positive beliefs about RA will increase RQ effect on intention to use RA	Ricci, Rokach and Shapira (2022), Xiao and Benbasat (2007), Whang and Im (2018)
Content-based filtering	Generates recommendations based on users' interests and past purchase behavior	Recommends niche and personalized items specific to each user	Limited ability to expand on users' existing interests	Lack of ability to recommend products beyond consumers' preferences weakens RQ effect on decision-making satisfaction	Despite receiving personalized recommendations, lack of exposure to new products weakens RQ effect on RA satisfaction	Not receiving novel recommendations will lead to unclear RQ effect on consumer intention to use RA	Jiang and Benbasat (2004), Köcher et al. (2019), Ricci, Rokach and Shapira (2022)
Interactive RA	Generates "in-the-moment," "to-the-point" recommendations	More widespread, time-efficient, and pragmatic recommendations	More complex process and difficult for user to use and understand mechanism	Consumers' evaluation of RA is complex and beyond RQ, making RQ effect on decision-making satisfaction unclear	Despite working with dynamic and interactive RA, complexity of using it makes RQ effect on RA satisfaction unclear	Receiving pragmatic and time-efficient recommendations will enhance RQ effect on intention to use RA	Van Doorn et al. (2017), Wirtz et al. (2018)
Self-serving recommendation	Generates featured and straightforward recommendations based on firm's interest	Higher sales and revenue for firms	Recommendation is not according to consumers' best interests	Straightforward recommendations enhance RQ effect by simplifying decision making	Less trust in RA intention weakens RQ effect on RA satisfaction	Receiving straightforward but unhelpful recommendations makes effect of RQ on intention to use RA unclear	Hunold, Kesler and Laitenberger (2020), Whang and Im (2018)
RA Presentation							
Solicited recommendation	Generates recommendations based on consumer's direct request	More relevant and helpful recommendations	Higher consumer effort in the process	More accurate recommendations enhance RQ effect on decision-making satisfaction	Despite receiving personalized recommendations, higher effort by consumers makes effect of RQ on RA satisfaction unclear	Higher effort by consumers and less credit to RA weakens RQ effect on intention to use RA	Marchand and Marx (2020), Tsekouras, Li and Benbasat (2022)
Comprehensiveness of results	Generates useful list of recommendations	Broader exposure to many different products	More consumer effort in decision making	Despite receiving useful list of items, consumers experience more difficulty in decision making, making effect of RQ on decision-making satisfaction unclear	Consumers being overwhelmed with many recommendations makes effect of RQ on RA satisfaction unclear	Despite being exposed many varied products, consumers may be overwhelmed, making the effect of RQ on intention to use RA unclear	Huang and Zhou (2019), Xiao and Benbasat (2007)
Website-embedded RA	RAs are directly embedded in retailers' websites	Everything consumers need is in same place	Other factors (e.g., website quality) affect consumers' evaluation of the RA	Despite easy access, more complicated decision-making makes effect of RQ on decision-making satisfaction unclear	More factors affecting RA evaluation enhances ambiguity, making effect of RQ on RA satisfaction unclear	Difficulty in evaluating RA performance makes effect of RA on intention to use RA unclear	Nilashi et al. (2016), Whang and Im (2021)

(*continued on next page*)

Table 1 (*continued*)

RA characteristics	Mechanism	Advantages	Disadvantages	Effect on decision-making satisfaction	Effect on RA satisfaction	Effect on future use intention	Sources
RA Data							
Location-based RA	Generates recommendations using consumer location information	More relevant recommendations based on consumers' particular preferences	Only applies to location-sensitive choice	Recommendations based on one piece of information (location), which makes effect of RQ on decision-making satisfaction unclear	Recommendations do not capture all aspects of consumer preferences, leading to a negative effect of RQ on RA satisfaction	Improving user experience by reducing number of searches available enhances RQ effect on intention to use RA	Divyaa and Pervin (2019), Rikitianskii, Harvey and Crestani (2014), Zhu et al. (2014)
Social network–based RA	Generates recommendations using consumer social information	Easy access to other users with similar preferences (i.e., people in consumers' network)	Privacy concerns regarding access to sensitive information	Privacy concerns make effect of RQ on decision-making satisfaction unclear	Ability to provide socially acceptable recommendations might enhance RQ effects on RA satisfaction	Less trust in RA intention and higher privacy concerns weakens RQ effect on intention to use RA	Chen et al. (2019), Ricci, Rokach and Shapira (2022), Zhu et al. (2014)
Direct input from consumers	Consumers participate in co-producing recommendations by providing direct information	Better fit, more accurate recommendations	Greater consumer effort in the process	Increased consumer effort needed makes effect of RQ on decision-making satisfaction unclear	Increased consumer effort needed makes effect of RQ on satisfaction with RA unclear	Increased consumer effort needed and less credit given to RA weakens RQ effect on intention to use RA	Bendapudi and Leone (2003), Tsekouras, Li and Benbasat (2022)
Personalized recommendation	Generates personalized recommendations using consumers' personal data	More personalized, accurate recommendations	Not easy to access all of consumers' personal data	Greater focus on personal data makes effect of RQ on decision-making satisfaction unclear	Despite accuracy of recommendations, increased input from consumers needed makes effect of RQ on RA satisfaction unclear	Higher effort and input from consumers makes effect of RA on intention to use RA unclear	Senecal and Nantel (2004), Whang and Im (2018)

novelty reflects the degree to which the RA can assist consumers in discovering new items. If RAs only suggest items similar to consumers' expressed preferences, their recommendation novelty may appear limited (Castells, Hurley and Vargas 2022). If they seem overly similar (i.e., low diversity), suggestions also might have limited value or be perceived as biased (Nilashi et al. 2016). Perceived novelty and diversity are both desirable to consumers, who tend to engage in variety-seeking behavior. These operationalization differences have two implications for our meta-analysis. First, we include all empirical studies, regardless of construct operationalizations (i.e., accuracy, novelty, and diversity). Second, we control for whether our results differ depending on the operationalization used.

Classification of RA types

Tsekouras, Li and Benbasat (2022, p. 3) explain that "RAs vary based on the sources of information they use, their decision strategies, and the degree of interaction between consumers and firms in crafting the recommendations." Accordingly, to assess performance differences, we classify RAs in terms of the underlying algorithm, recommendation presentation, and data source, as we summarize in Table 1.

RA algorithm. To establish recommendations, RAs apply several types of algorithms. Two popular approaches are *content-based filtering* and *collaborative filtering*, which indicate whether the algorithm relies on consumers' own stated preferences and past purchase behavior or decisions made by similar consumers (Xiao and Benbasat 2007). Content-based filtering tends to recommend "content" (i.e., products) similar to what a consumer has liked or purchased in the past, such that these recommendations are specific to the consumer. Collaborative filtering instead might expand the consumer's existing interests and recommend a wider range of products (Ricci, Rokach and Shapira 2022). More recently, scholars examined *interactive RAs* that rely on AI technology—a class of machine learning algorithm used to predict and suggest products to consumers. AI techniques have been used extensively to enhance decision-making quality in many areas (Schuetzler, Grimes and Giboney 2020). Interactive RAs help consumers overcome information overload, improve the level of personalization (Liao, Widowati and Chang 2021), and propose more widespread, time-efficient, and pragmatic recommendations (Anastasiya 2021, Shi, Gong and Gursoy 2021). Whereas traditional algorithms rely on consumers' own purchase history, interactive RAs use AI to analyze their interactions and make so-called in-the-moment, to-the-point recommendations. Furthermore, some algorithms, known as *self-serving recommendations* (e.g., "Amazon's choice"), make recommendations that benefit the firms over consumers. Self-serving recommendations can generate featured and straightforward recommendations that simplify decisions for consumers but are not always based on consumers' best interest (Hunold, Kesler and Laitenberger 2020).

We examine the effects of RAs that use these various algorithms in our meta-analysis. Extant studies examine the effect of perceived recommendation quality, but few have tested whether its effects on outcomes (e.g., future use intentions) vary across algorithms, with the exception of studies of collaborative versus content-based filtering (see Web Appendix A).

RA presentation. Recommendations get presented to consumers in various ways. Some RAs present *solicited recommendations* in response to consumers' requests: When consumers look for a specific product, they indicate their preferences to the RA to generate recommendations (Tsekouras, Li and Benbasat 2022). Consumers are more involved in such a process. Other RAs provide recommendations even before consumers have requested them directly (Marchand and Marx 2020), which might be based on consumers' purchase history and decision process. Although solicited recommendations may be more relevant, they require greater user effort than do unsolicited recommendations. In addition, RAs differ in *comprehensiveness*; some recommend limited options, and others suggest a long list of products. The latter case could provide broader exposure to a variety of items, but because the RA's purpose is to facilitate consumers' decision-making and help them cope with information overload, such a comprehensive list might be problematic, in that it increases decision-making difficulty (Goodman et al. 2013). Finally, some RAs are *embedded directly in the retailer's website* (e.g., Walmart.com), so recommendations appear at the moment consumers make purchases, which makes it easier for consumers to access everything in one place. Other RAs might use a separate platform, such as mobile applications, to present product recommendations. Extant research has not tested the influence of RA presentation on the effects of perceived recommendation quality on different outcomes (Web Appendix A).

RA data. RAs can use different data sources to provide recommendations. *Location-based RAs* use consumers' geographic information (Divyaa and Pervin 2019, Zhu et al. 2014); they recognize and consider the unique features of each consumer's current location to make recommendations. However, these recommendations cannot capture all aspects of consumer preferences. Another important data source that some RAs have integrated is *social media*. Information about consumers and their social networks may improve understanding of their preferences and enhance prediction quality (Ricci, Rokach and Shapira 2022). However, privacy concerns arise related to access to sensitive consumer information (Zhu et al. 2014). Moreover, to gather consumer data, some RAs use platforms that ask consumers directly to provide input about their preferences; others obtain consumers' historical data to predict their preferences. Receiving *direct input from consumers* may enable RAs to provide more personalized recommendations that fit their needs (Xiao and Benbasat 2007) but also may undermine the RA's importance for consumers, who perceive it as less efficient (Tsekouras and Li 2015). Generally, consumers' personal data may be used to provide recommendations, such that many RA types (e.g., collaborative, content-based, location-based filtering) can be broadly classified as *personalized RAs*, in the sense that they use personal data. Their differences pertain to which specific data are used and how they are processed by different algorithms. Nonpersonalized RAs do not leverage any personal data and provide more generic recommendations, such as suggesting top-selling products (Ghiassaleh, Kocher and Czellar 2020, Ricci, Rokach and Shapira 2022). To the best of our knowledge, no research has assessed whether perceived recommendation quality effects differ depending on RA data sources (see Web Appendix A).

In summary, several studies examine single recommendation techniques, but a comprehensive performance assessment of different RA types is lacking. We classify RAs according to their algorithm, recommendation presentation, and data source, and we assess whether perceived recommendation quality effects vary with these characteristics. Our assessment provides insights into the performance differences achieved by various RA types and also suggests guidance for continued research into promising RAs.

Meta-analytic framework

Building on Xiao and Benbasat's (2007) conceptual study, Fig. 1 depicts our guiding meta-analytic framework. It comprises (1) the direct effect of perceived recommendation quality on future use intention, (2) the mediating effects of decision-making satisfaction and RA satisfaction, and (3) the moderating effects of RA types. Table 2 provides construct definitions.

First, we include perceived recommendation quality as the focal construct, along with its direct impact on consumers' intentions to use the RA in the future (distal outcome of RA use), as the central relationship of interest (Castells, Hurley and Vargas 2022).

Second, we select decision-making and RA satisfaction (proximal outcomes of RA use) as mediators. This choice is consistent with Xiao and Benbasat's (2007) identification of two groups of mediators (consumer decision-making and consumer evaluation of RAs). For parsimony, we include satisfaction as a broad construct that captures various aspects of consumer decision-making and evaluation of RAs, which aligns with prior decision-making and information systems research (Lee and Choi 2017, Pu, Chen and Hu 2011). Accordingly, "decision-making satisfaction" refers to consumers' satisfaction with the decision-making process and outcome,

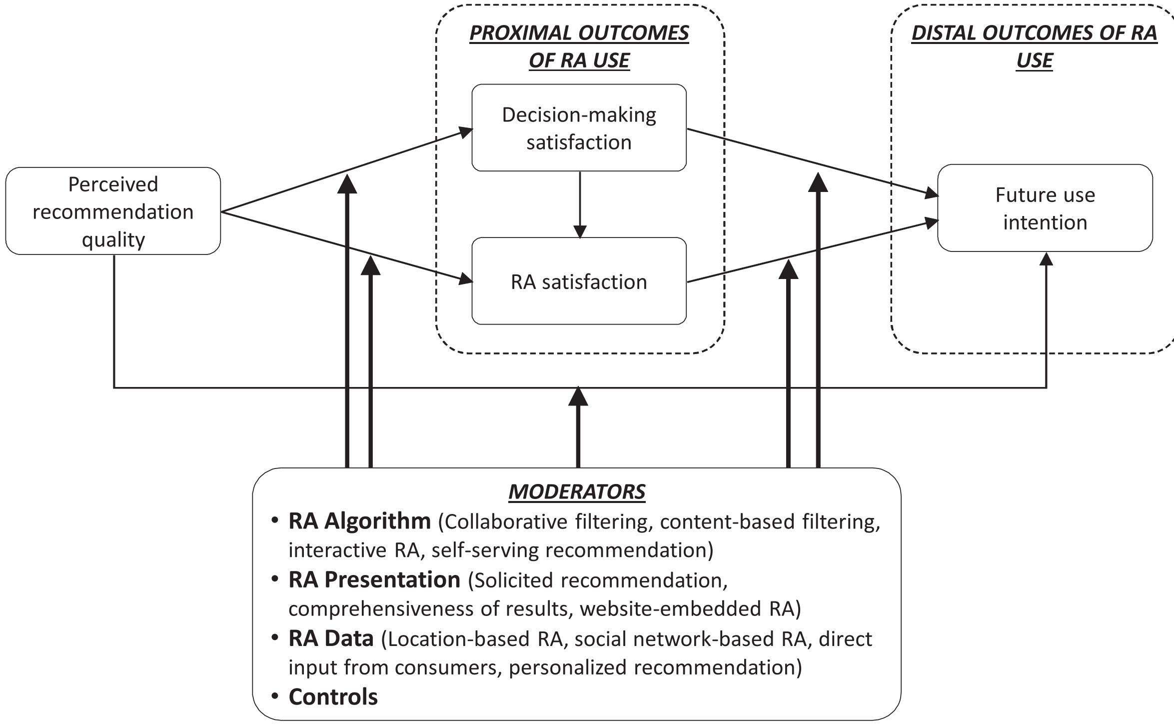

Fig. 1. Meta-analytic framework testing the performance of RA types.

and it encompasses consumers' choice confidence, choice satisfaction, decision effort, and decision quality (Xiao and Benbasat 2007); "RA satisfaction" refers to consumers' satisfaction with an RA in terms of its perceived usefulness, ease of use, and trustworthiness (Xiao and Benbasat 2007). By adding these two mediators to the quality–satisfaction–intention chain in our framework, we ensure it is well-grounded in marketing literature (Frank et al. 2014) and supported by prior RA studies (Yoon et al. 2013). It is also consistent with the information systems success model, which suggests that consumers' quality perception is a primary influence on their satisfaction (DeLone and McLean 1992).

Third, we address the moderating roles of RA types in terms of the underlying algorithm, recommendation presentation, and data source. This focus aligns with context-specific theorizing in information systems research (Hong et al. 2014), in which technology types (characteristics of the technology artifact) as contextual factors function as potential moderators of the relationships in technology acceptance models (Schepers and Wetzels 2007). For example, Verma et al. (2023) assess the moderating role of platform type on the effects of electronic word of mouth. Even if RA types might influence perceived recommendation quality directly (Yoon et al. 2013), we consider their role as context-specific moderators.

We first explore whether RA types moderate the effects of perceived recommendation quality on decision-making satisfaction, RA satisfaction, and future use intention. According to information diagnosticity theory, perceived recommendation quality is an important piece of information that consumers use to evaluate an RA's performance, and its impact varies with its perceived diagnosticity (Khare, Labrecque and Asare 2011). We expect differences in the diagnosticity of perceived recommendation quality across RA types, because they use different algorithms and data to generate recommendations and present them in different ways. Moreover, we explore whether RA types moderate the impact of the two satisfaction mediators on future use intention. Marketing literature suggests that the effects of consumer satisfaction vary with various context factors (Seiders et al. 2005).

For this investigation, we adopt an empirics-first approach that "(1) is grounded in (originates from) a real-world marketing phenomenon, problem, or observation, (2) involves obtaining and analyzing data, and (3) produces valid marketing-relevant insights without necessarily developing or testing theory" (Golder et al. 2022, p. 1). Thus, rather than developing specific hypotheses, we use our comprehensive data set to conduct a meta-analysis of all possible effects in the framework and estimate their effect sizes. This approach is more apt for our study than a dominant theory-first approach, for two reasons. First, the empirics-first approach works well with meta-analyses (Datta et al. 2022). Second, we could formulate hypotheses for main and mediating effects, but most of the RA type moderators are novel, and their moderating effects have not been examined before. An empirics-first approach allows us to use our rich data set to probe these relationships with an open mind and thus discover new insights relevant to real-world retailers using RAs (Golder et al. 2022).

Table 2
Constructs, definitions, aliases, and examples.

Main Variables	Definition	Alias(es) / Examples
Perceived recommendation quality	The degree to which the RA can assist consumers in decision making; includes consumers' perception of product recommendations as accurate, novel, and diverse (Nilashi et al. 2016)	Advice quality, perceived fit of recommendation, perceived accuracy, perceived diversity, perceived informativeness, novelty, perceived personalization, perceived quality, perceived relevance, and perceived variety
Decision-making satisfaction	Consumers' satisfaction with the decision-making process and outcomes. Includes consumer's choice confidence, choice satisfaction, decision effort, and decision quality (Xiao and Benbasat 2007)	Choice confidence, choice satisfaction, decision effort, decision satisfaction, decision time, decision quality, decision-making quality, satisfaction with outcome
RA satisfaction	Consumers' satisfaction with an RA in terms of perceived usefulness, ease of use, and trustworthiness (Xiao and Benbasat 2007)	Ease of use, effort expectancy, helpfulness, performance expectancy, usefulness, trust (benevolence, competence, integrity)
Future use intention	The strength of a person's intention to perform a specified behavior (Fishbein and Ajzen 1975)	Intention to use, willingness to use, adoption intention
Moderators		
RA Algorithm		
Collaborative filtering	RAs that generate recommendations based on the interests of similar consumers	Dummy-coded 1 when an RA uses similar consumers' past product purchases or ratings (e.g., MovieLens), and 0 otherwise.
Content-based filtering	RAs that generate recommendations based on the consumer's own stated preferences and past purchase behavior	Dummy-coded 1 when an RA uses a consumer's own product preference or purchase history, and 0 otherwise.
Interactive RA	RAs that use AI to analyze user interactions and make "in-the-moment," "to-the-point" recommendations	Dummy-coded 1 when an RA is conversational and interacts with a consumer (e.g., Siri), and 0 otherwise.
Self-serving recommendation	RAs that feature recommendations based on firm's interest	Dummy-coded 1 when an RA serves the interest of a firm/retailer in recommendation (e.g., "Amazon's choice"), and 0 otherwise.
RA Presentation		
Solicited recommendation	RAs that provide recommendations based on consumer request	Dummy-coded 1 when an RA presents recommendations only upon a consumer's explicit request, and 0 otherwise.
Comprehensiveness of results	RAs that recommend a useful list of items	Dummy-coded 1 when an RA recommends a larger number of items to a consumer, and 0 otherwise.
Website-embedded RA	RAs that are directly embedded in retailers' websites	Dummy-coded 1 when an RA presents recommendations within a retailer's website (e.g., Walmart.com), and 0 otherwise.
RA Data		
Location-based RA	RAs that use consumer location information to provide recommendations	Dummy-coded 1 when an RA uses a consumer's explicit location information such as ZIP code and real-time GPS data, and 0 otherwise.
Social network–based RA	RAs that use consumer social information to provide recommendations	Dummy-coded 1 when an RA uses a consumer's social data (e.g., social network friends) from his or her social media, and 0 otherwise.
Direct input from consumers	RAs that make consumers participate in co-producing recommendations by providing direct information	Dummy-coded 1 when an RA requests and uses a consumer's onsite direct input (e.g., search terms or criteria), and 0 otherwise.
Personalized recommendation	RAs that use personal consumer data to make personalized recommendations	Dummy-coded 1 when an RA uses a consumer's various personal data, such as location, profile information, browsing history, and past purchase, and 0 otherwise.

Main and mediating effects

Some studies indicate that perceived recommendation quality is related to consumers' intentions to use the RA in the future (Aggarwal and Vaidyanathan 2005), often in reference to the technology acceptance model and with the argument that consumers' beliefs about tools such as RAs influence their intentions to use them (Venkatesh et al. 2003). Regarding mediating mechanisms, scholars suggest that perceived recommendation quality influences consumers' decision-making satisfaction in terms of their choice confidence, choice satisfaction, decision effort, and decision quality (Xiao and Benbasat 2007, Yoon et al. 2013). Some scholars also argue that perceived recommendation quality influences RA satisfaction (Hess, Fuller and Campbell 2009). Referring to the information systems success model, these studies contend that perceived recommendation quality is a technology belief that improves consumers' satisfaction with the RA in terms of its usefulness, ease of use, and trustworthiness (DeLone and McLean 1992, Xiao and Benbasat 2007).

RA algorithm as moderator

Collaborative filtering. Collaborative filtering involves generating consumer recommendations based on the likes and dislikes of similar consumers (Ricci, Rokach and Shapira 2022). This popular technique can provide more credible and diagnostic suggestions for consumers, which is especially helpful when consumers are unfamiliar with a product category (Xiao and Benbasat 2007). Consumers may find the recommendations relevant to them, which should then increase their confidence in the RA (Whang and Im 2018). Thus, this algorithm might strengthen the impact of consumers' perceived recommendation quality on their decision-making satisfaction and future use intentions. In addition, easier interactions with and higher confidence in collaborative-filtering RAs likely result in a stronger effect of perceived recommendation quality on RA satisfaction. The popularity of and consumer familiarity with this algorithm could enhance the positive effects of the two satisfaction mediators on future use intentions too.

Content-based filtering. The content-based filtering technique, which is based on consumers' own stated preferences for product attributes and their past purchase behavior (Ricci, Rokach and Shapira 2022), is chiefly known for the relevance of the recommendations it produces. On the one hand the recommendation thus should appear straightforward and transparent, and consumers are likely more certain in their evaluation of its quality. Moreover, the recommendation is attribute based, so consumers may find the information more diagnostic and helpful, leading to more effective and efficient choice decisions (Jiang and Benbasat 2004, Köcher et al. 2019). On the other hand, because content-based filtering cannot recommend new products beyond consumers' own preferences (Ricci, Rokach and Shapira 2022), it could weaken the effects of perceived recommendation quality for some outcomes. The overall moderating effect of this algorithm is unclear. Similarly, whether and how this algorithm moderates the effects of the satisfaction mediators on future use intentions is unclear, because recommendations tend to be high in relevance but low in novelty and diversity.

Interactive RA. A typical RA is an invisible software system that produces text- or picture-based "single-shot" product recommendations. According to the concept of perceptual salience (Bhattacharyya et al. 2022), consumers' evaluation of such traditional RAs tends to focus on recommendation quality, because no real consumer–RA interactions occur; the recommendation is the only visible, diagnostic element consumers have. In contrast, interactive RAs such as robo-advisors and chatbots are dynamic, and consumers' use and evaluations thereof may be more complex, influenced by social, emotional, and relational factors (Van Doorn et al. 2017, Wirtz et al. 2018). With interactive RAs, perceived recommendation quality might be less salient for determining decision-making satisfaction and RA satisfaction. Conversely, the effect of perceived recommendation quality on future use intentions might be enhanced, because interactive RAs are fun to use and provide more pragmatic, time-efficient recommendations. In terms of the effects of both satisfaction mediators on future use intentions, interactive RAs might be fun to use, or using them could feel difficult and complex. Fun and enjoyment could strengthen these links, whereas complexity and inconvenience could weaken them.

Self-serving recommendation. Some RAs generate "featured" or "top pick" recommendations that benefit firms/retailers over consumers (Hunold, Kesler and Laitenberger 2020). Firms often pay for these self-serving or biased recommendations to maximize revenue rather than help consumers make the best choice (Xiao and Benbasat 2018). Therefore, even if consumers deem such recommendations' quality acceptable, they may have reservations, not fully believe the RA (Whang and Im 2018), and express lower RA satisfaction. However, self-serving recommendations like "top picks" also might enhance the effect of perceived recommendation quality on decision-making satisfaction, because they are straightforward and can serve as a heuristic to facilitate and simplify consumer decision-making. Because this algorithm could have opposite effects on the two satisfaction mediators, predicting how it moderates the impact of perceived recommendation quality is difficult. But due to decreased confidence in the RA's motivation, we expect both satisfaction mediators to have weaker impacts on future use intention in the case of self-serving recommendations.

RA presentation as moderator

Solicited recommendation. In e-commerce, RAs can make product recommendations with or without consumer requests. Some provide recommendations only when sought, and others do so automatically, without consumers asking for or expecting it (Marchand and Marx 2020). Compared with automated or unsolicited recommendations, the quality of solicited recommendations should be higher in terms of relevance and helpfulness, which might amplify the effect of perceived recommendation quality on decision-making satisfaction. Conversely, consumers' perceived RA effort for solicited recommendations tends to be lower (Tsekouras, Li and Benbasat 2022) because, according to attribution theory and self-serving bias (Bendapudi and Leone 2003), consumers typically attribute recommendation quality more to their own effort. Thus, perceived recommendation quality could be less relevant in determining RA satisfaction and predicting future use intentions. Similarly, if consumers take credit for their choices, decision-making satisfaction may be less relevant in predicting future use intention.

Comprehensiveness of results. As noted, RAs offer recommendation lists with varying length, such that some present limited options, but others offer a longer list. Consumers faced with more recommended options have more choices but also must exert more decision-making effort (Xiao and Benbasat 2007). Information diagnosticity also does not increase with information quantity (Filieri 2015). When consumers receive fewer recommendations, they have fewer choices, but their perceived information diagnosticity tends to be higher, such that they perceive less decision effort and choice difficulty (Huang and Zhou 2019). Thus, it is unclear whether RA comprehensiveness strengthens or weakens the effects of perceived recommendation quality on decision-making satisfaction, RA satisfaction, and future use intention; it might even reveal a complex, curvilinear moderating effect (e.g., inverted-U relationship: too few are limiting and too many are overwhelming), which we consider in our analyses. As comprehensiveness increases, inconvenience in consumer decision-making and RA use inevitably increases, so we anticipate a weaker effect of the two satisfaction mediators on future use intentions.

Website-embedded RA. Consumers can find RAs in a variety of environments. Website-based RAs are embedded directly in a retailer's website (e.g., Amazon, Walmart) (Xiao and Benbasat 2007); with the growing popularity of other platforms such as mobile devices and apps, they also increasingly appear as app-based RAs (e.g., Alexa, Siri). Consumers perceive recommendations presented on retailers' websites versus other platforms differently (Whang and Im 2021). They tend to regard nonwebsite platforms as the RA itself and evaluate them according to the recommendation content (i.e., recommendation quality), which directly influences their decision-making satisfaction, RA satisfaction, and future use intentions. Conversely, consumers tend to consider a website-embedded RA as part of the retailer's website, so their evaluation and acceptance is affected by other factors, such as website quality (Nilashi et al. 2016). Thus, perceived recommendation quality might carry less weight in determining decision-making satisfaction and RA satisfaction or predicting future use intentions. However, the website context seems unlikely to have any influence on the effects of the two satisfaction mediators on future use intentions though.

RA data as moderator

Location-based RA. Location-based RAs incorporate real-time consumer location information (prominent, explicit data) to provide recommendations (Divyaa and Pervin 2019, Zhu et al. 2014). These RAs extract extensive knowledge about particular consumers' preferences and behavior and accordingly provide relevant recommendations (Rikitianskii, Harvey and Crestani 2014). When an RA relies on explicit and current information, such as location, the user experience should be improved, due to reduced search effort, and the effect of perceived recommendation quality on consumers' future use intentions might be stronger. However, the influence of such RAs on the two satisfaction mediators is unclear. It could have a negative moderating effect, because location-based RAs rely on just one piece of information and may not generate the best choices for consumers. Moreover, due to limited capabilities of such RAs, the effects of both satisfaction mediators on future use intentions might be weakened. Alternatively, the convenience of such recommendations could enhance these relationships.

Social network–based RA. With the rise of social networks (e.g., Twitter, Facebook), RAs began to tap into consumers' social networks to extract implicit information (e.g., interactions with family, friends, colleagues) and use it to make recommendations (Chen et al. 2019, Ricci, Rokach and Shapira 2022). Social data include personal information that is highly sensitive and private. Consumers may sense their privacy is being invaded and threatened by RAs if they receive social network–based recommendations (Zhu et al. 2014). Thus, even if recommendation quality seems high, consumers may choose not to use these recommendations, their privacy concerns outweigh their helpfulness. However, social network–based RAs also can provide socially acceptable recommendations, which might enhance the effects of perceived recommendation quality on consumers' decision-making satisfaction and RA satisfaction. Due to information sensitivity and privacy concerns, even if consumers are satisfied with a social network–based RA and their decision-making, their future use intentions still might be lower.

Direct input from consumers. Whereas some RAs use backstage consumer data, others make recommendations by requesting direct input from consumers (Xiao and Benbasat 2007). For example, online travel agents ask consumers to provide search terms or criteria for the system to find matches. Consumers may be happy to provide input for their own benefit, though they also could perceive the RA's effort as lower and their own effort as higher in this process (Tsekouras, Li and Benbasat 2022). Because they co-produce the recommendation, they likely give themselves more credit than the RA for the quality of the results (Bendapudi and Leone 2003). Thus, perceived recommendation quality might be less relevant and diagnostic in determining consumers' decision-making satisfaction and RA satisfaction or predicting their future use intentions. Similarly, consumers tend to attribute their choices to their own, rather than to the RA's effort, so decision-making satisfaction could be less relevant in predicting their future use intentions.

Personalized recommendation. Recommendations differ in their degree of personalization. Most RAs in e-commerce use various personal consumer data, including profile information, browsing history, and past purchases, to make personalized recommendations, but some provide nonpersonalized recommendations, such as "top 10 selections" (Ricci, Rokach and Shapira 2022). The former data generate more relevant and helpful product suggestions, by considering each consumer's unique characteristics, interests, and preferences; the latter tend to ignore individual differences and make recommendations that are not correlated with the consumer's needs (Senecal and Nantel 2004). Therefore, we posit that personalized recommendations offer greater information diagnosticity (Whang and Im 2018), which can amplify the effects of perceived recommendation quality on decision-making satisfaction, RA satisfaction, and future use intention. On the other hand, personalized recommendations use personal data or require consumer input, so privacy concerns and consumer effort perceptions might weaken these effects. Thus, the moderating role of personalized recommendations is unclear. Similarly, we cannot predict whether the effects of the satisfaction mediators on future use intentions differ between personalized and nonpersonalized recommendations.

Controls

Extant studies contain various characteristics; we control for moderators that describe the type of product and market examined (search goods, goods vs. services, product knowledge, and risk), study design (real RA interaction, experiment, and endogeneity concerns), sampling approach (student sample, sample age, sample gender, multiple industries, and U.S. vs. non-U.S.), and perceived recommendation quality operationalization (accuracy vs. others). Furthermore, we control for publication outlet (publication status and marketing vs. nonmarketing journal) and time effects (study year).

Method

Study search and inclusion criteria

We searched electronic databases, such as EBSCO and ResearchGate, to identify studies for the meta-analysis. We used different keywords, including “recommendation agent,” “recommendation systems,” “recommender systems,” “automated recommendations,” “online recommendations,” and “product recommendations,” as well as alternative wordings. We also leveraged search terms that reflect advanced recommendation technology, such as “advisory systems,” “artificial intelligence,” “AI,” “IBM Watson,” “Soul Machines,” “robot,” “robo-advisor,” “chatbots,” “virtual assistance,” and “voice assistant,” in combination with “recommendation” or “recommender.” We searched for these keywords in Google Scholar, which helped us identify further studies, unpublished studies, studies in conference proceedings, and dissertations. We also examined the collected studies’ references and citations of key papers in the field (e.g., Xiao and Benbasat 2007).

We applied three inclusion criteria: Studies had to (1) examine RAs from the consumer’s perspective, as defined in the present study (i.e., we excluded studies that examined other technologies or online recommendations/reviews by other consumers); (2) measure any two constructs of the meta-analytic framework; and (3) report correlations or other statistical information that could be converted to correlations (i.e., we excluded qualitative studies and conceptual papers). The meta-analysis thus featured 98 studies (see Web Appendix B).

Effect size

We used the correlation coefficient as the effect size, because most studies in our meta-analysis reported correlations; moreover, this effect size is comparable across collected studies (Grewal, Puccinelli and Monroe 2018). Around 85% of meta-analyses published in major marketing journals use this effect size (see Table 2 in Grewal, Puccinelli and Monroe 2018). Because some studies use experiments, we also converted other statistical information (t-values, means, and SDs) to correlations. We did not convert regression weights to correlations, a practice that has prompted some criticism of being associated with potential downward biases for calculating the integrated effect sizes (Roth et al. 2018). We averaged the effect sizes if one independent sample reported more than one correlation for the same relationship, which is necessary to avoid giving one independent sample excessive weight in the subsequent analyses (Grewal, Puccinelli and Monroe 2018). The final data set, after averaging, includes 480 correlations reported in 122 independent samples by 98 studies. The cumulative sample size is 32,172.

Coding

All three authors extracted information from the collected studies about effect sizes and other statistical information. They are experts in the field, hold doctorates, and are familiar with relevant theory and definitions. Two authors coded the same set of studies, and the third checked the quality of coding and conducted the analyses. We classified the effect sizes according to the construct definitions displayed in Table 2. The interrater reliability was 96%, and the author team resolve any differences through discussions. We extracted information about the reliabilities of constructs, sample sizes, and other study characteristics that may function as moderators. Similar to other meta-analyses (Blut et al. 2021), we used dummy coding for moderators; most RA characteristics are categorical variables, and this approach eases comparability. Specifically, we dummy-coded the RA algorithm, including collaborative filtering (1 = yes, 0 = no), content-based filtering (1 = yes, 0 = no), interactive RA (1 = yes, 0 = no), and self-serving recommendation (1 = yes, 0 = no). We next coded the moderators describing RA presentation: solicited recommendation (1 = yes, 0 = no), comprehensiveness of results (1 = many, 0 = few), and website-embedded RA (1 = yes, 0 = no). Then, we coded the RA data, including location-based RA (1 = yes, 0 = no), social network–based RA (1 = yes, 0 = no), direct input from consumers (1 = yes, 0 = no), and personalized recommendation (1 = yes, 0 = no). Finally, we extracted further information to control for the influence of other study characteristics (see Web Appendix E).

Descriptive analyses

We used Hunter and Schmidt’s (2004) approach to meta-analysis, a random-effects approach that corrects the effect sizes for various artifacts. Specifically, we first corrected the correlation coefficients for measurement unreliability in the dependent and independent variables. We then weighted the reliability-corrected correlations by sample size to correct for sampling error. We complemented Hunter and Schmidt’s (2004) formulas for calculating 95% confidence intervals around the reliability-adjusted and sample-size–weighted correlations with additional power tests of statistical analyses. Hunter and Schmidt (2004) further suggest estimating 80% credibility intervals, which give indications of the variance in effect sizes. As recommended by

Table 3
Descriptive results on consequences of perceived recommendation quality.

Relationship	k	N	Simple aver. correlation (r)	Sample-weighted, rel.-adj. correlation (rwc)	CI_{95-}	CI_{95+}	CR_{80-}	CR_{80+}	Q	FSN	Power
RQ → Decision-making satisfaction	21	4799	.33	.41*	.31	.51	.12	.70	202*	3713	>.999
RQ → RA satisfaction	43	9903	.45	.49*	.42	.57	.19	.80	450*	26335	>.999
RQ → Future use intention	39	8682	.48	.49*	.41	.57	.18	.81	425*	22459	>.999
Decision-making satisfaction → Future use intention	23	5283	.30	.42*	.29	.54	.04	.79	377*	4619	>.999
RA satisfaction → Future use intention	63	16928	.52	.59*	.54	.65	.33	.85	582*	92527	>.999
Decision-making satisfaction → RA satisfaction	16	3314	.30	.46*	.35	.58	.17	.75	145*	2167	>.999

RQ = recommendation quality, k = number of effect sizes, N = cumulative sample size, CI = 95% confidence interval, CR = 80% credibility interval, Q = Q statistic of heterogeneity, FSN = fail-safe N, Power = results of power test. * $p < .05$.

meta-analysis best practices (Grewal, Puccinelli and Monroe 2018), we estimated Q-tests of homogeneity as additional tests for effect size variance (some early meta-analyses suggest this test to determine whether a moderator analysis is needed; however, current best practice is to conduct moderator tests independent of the test results). Then, we calculated Rosenthal's (1979) fail-safe N (FSN) as suggested by Grewal, Puccinelli and Monroe (2018), which indicates the robustness of the results and the influence of publication bias. The FSN identifies the number of studies with null results that would be needed to lower a significant effect size to a barely significant level ($p = .05$). Rosenthal (1979) suggests that results are robust when FSNs are greater than 5 × k + 10, where k equals the number of correlations. Finally, we complemented publication bias tests with funnel plots; an asymmetric plot indicates potential publication bias.

Moderator tests

As Grewal, Puccinelli and Monroe (2018) suggest, we tested moderators in the meta-analysis with two approaches: subgroup analysis and meta-regression. Subgroup analysis indicates the strength and direction of effect sizes in different subgroups (e.g., collaborative filtering vs. other RAs). Then, we regressed the effect sizes on the moderators that were significant in the subgroup analysis. Meta-regression includes effect sizes (e.g., correlations between perceived recommendation quality and future use intentions) as the dependent variable and study characteristics (i.e., RA types and controls) as independent variables (Lipsey and Wilson, 2001). Borenstein et al. (2021) suggest using meta-regression to test the influence of multiple independent variables (moderators) at the same time to control for confounding variables.[2]

Results

Results of descriptive analyses

Table 3 displays the results of the descriptive analyses. All main and mediating effects in our framework are significant. The effect sizes are rather strong because we examine correlations among subjectively measured variables, which show some common method effects. The findings indicate the usefulness of the meta-analytic framework.

Perceived recommendation quality relates directly to the two mediators and consumers' intentions to use the RA in the future. Specifically, the effects of perceived recommendation quality on decision-making satisfaction (sample-weighted, reliability-adjusted average correlations [rwc] = .41, $p < .05$) and RA satisfaction (rwc = .49, $p < .05$) are significant, as are its effect on future use intentions (rwc = .49, $p < .05$). Moreover, the two mediators relate to consumers' intentions to use the RA in the future. Both decision-making (rwc = .42, $p < .05$) and RA (rwc = .59, $p < .05$) satisfaction display significant effects on future use intentions; the latter effect size is larger. Finally, we note that the two mediators are not independent, according to the significant effect of decision-making satisfaction on RA satisfaction (rwc = .46, $p < .05$). The results are similar to simple averaged correlations.

The Q-tests of homogeneity are significant in all cases, indicating substantial variance in effect sizes. This variance also is reflected in the wide credibility intervals and the forest plot in Fig. 2. The results do not indicate publication bias; all FSNs exceed the tolerance levels proposed by Rosenthal (1979). In addition, the funnel plots are symmetrical. Power tests suggest the sufficient power (>.5) of the analyses. When we test for the results of an effect size integration by perceived recommendation quality operationalization (Web Appendix C), the results remain similar for most relationships.

[2] The results of the subgroup and regression analyses are observational; they cannot prove causality (Borenstein et al. 2021). Thus, experimental studies are required to validate the causal nature of the moderator findings (Grewal, Puccinelli and Monroe 2018).

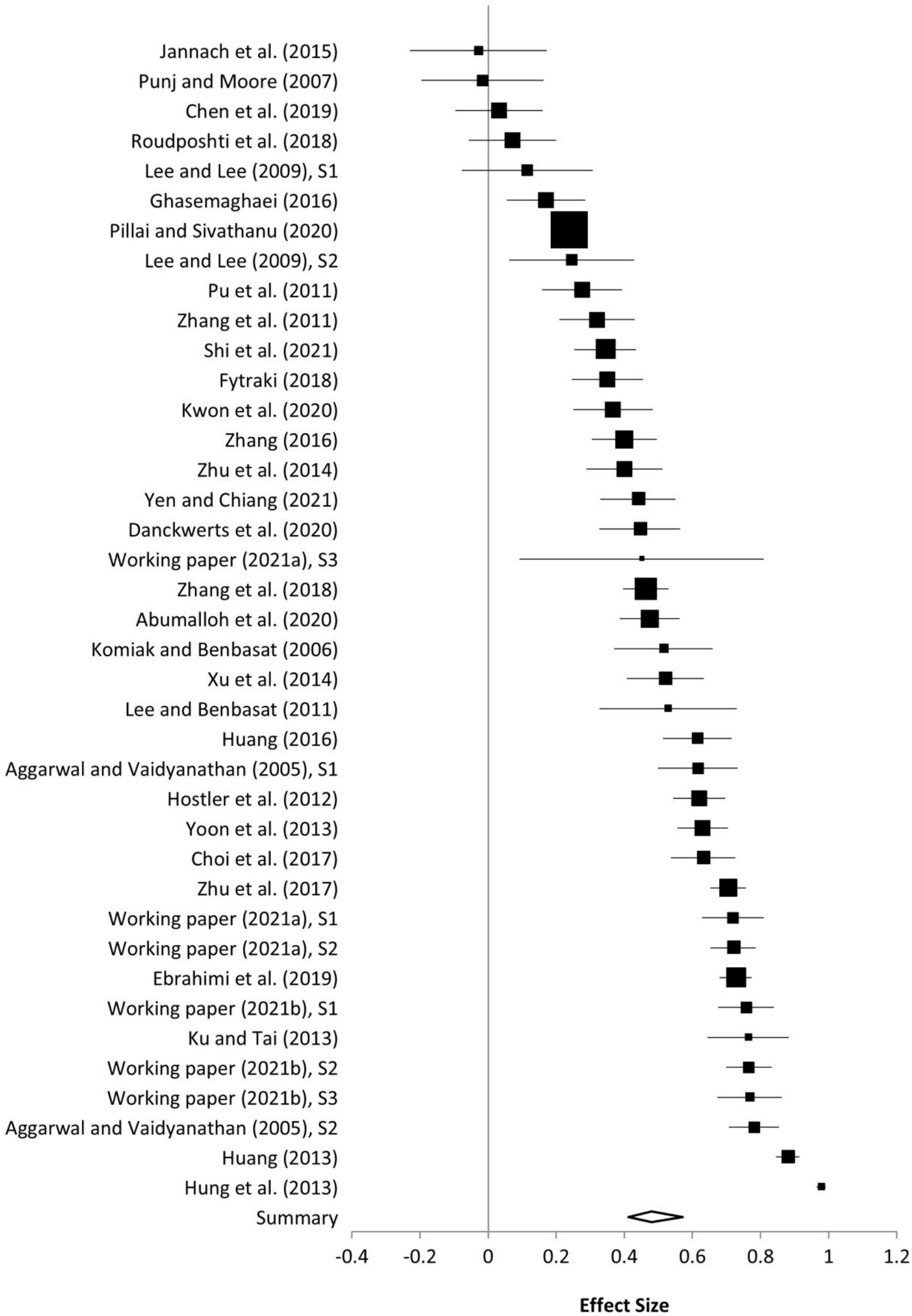

Fig. 2. Forest plot for relationship between perceived recommendation quality and future use intention (Abumalloh et al., 2020; Choi et al., 2017; Danckwerts et al., 2020; Ebrahimi et al., 2019; Fytraki, 2018; Ghasemaghaei, 2016; Hostler et al., 2012; Huang, 2013, 2016; Hung et al., 2013; Jannach et al., 2015; Komiak and Benbasat, 2006; Ku and Tai, 2013; Kwon et al., 2020; Lee and Benbasat, 2011; Lee and Lee, 2009; Pillai and Sivathanu, 2020; Punj and Moore, 2007; Roudposhti et al., 2018; Xu et al., 2014; Yen and Chiang, 2021; Zhang, 2016; Zhang et al., 2018; Zhu et al., 2017).

Results of moderator tests

We used a two-step approach to test the moderators. Grewal, Puccinelli and Monroe (2018, p. 22) suggest that "upon discovering significant differences between the various subgroups, the researchers should determine whether the difference persists after controlling for potential other moderators using meta-regression procedures." Thus, we first examine moderators using subgroup analysis. Table 4 displays the results. Multicollinearity is not a problem, as indicated by the low correlations among moderators (Web Appendix D). We also validated the initial moderator results by regressing the effect sizes on RA types and including control variables, for only those RA types that were significant in subgroup analysis. Similar to other meta-analyses, we included moderators if at least two observations per side of the moderator have been reported (Brown and Lam 2008; Maity, Dass and Malhotra 2014; Roschk, Loureiro and Breitsohl 2017). Several RAs enhance the effect of perceived recommendation quality on outcomes and the effects of the two mediators on future use intentions. However, some moderating effects turn nonsignificant when we control for other moderators. Again, multicollinearity is not a problem; the variance inflation factors range from 3.06 to 5.59. The results of significant moderators align with subgroup analysis (Table 5, Web Appendix E).

RA algorithm. We determine that *collaborative-filtering RAs* enhance the effects of perceived recommendation quality on RA satisfaction ($\beta = .23$, $p < .01$) and future use intentions ($\beta = .07$, $p < .05$). They also enhance the effect of decision-making satisfaction on future use intentions ($\beta = .23$, $p < .01$). *Content-based filtering RAs* weaken the effect of perceived recommendation quality on decision-making satisfaction ($\beta = -.17$, $p < .05$) and RA satisfaction ($\beta = -.10$, $p < .05$), as well as the effect of decision-making satisfaction on future use intentions ($\beta = -.24$, $p < .01$). We observe no differences for *interactive RAs*. *Self-serving recommendation RAs* enhance the effects of perceived recommendation quality on decision-making satisfaction ($\beta = .25$, $p < .05$) and future use intentions ($\beta = .11$, $p < .01$).

RA presentation. The results show that *solicited recommendation RAs* enhance the effect of perceived recommendation quality on decision-making satisfaction ($\beta = .34$, $p < .01$) but weaken the effects of both perceived recommendation quality ($\beta = -.26$, $p < .01$) and decision-making satisfaction ($\beta = -.23$, $p < .01$) on future use intentions. The *comprehensiveness of results* weakens the effect of RA satisfaction on future use intentions ($\beta = -.07$, $p < .01$).[3] We observe no differences for *website-embedded RAs*.

RA data. The *location-based RAs* enhance the effect of perceived recommendation quality on future use intentions ($\beta = .23$, $p < .01$) but weaken the effect of perceived recommendation quality on RA satisfaction ($\beta = -.10$, $p < .05$). *Social network–based RAs* enhance the effect of perceived recommendation quality on RA satisfaction ($\beta = .24$, $p < .01$) but weaken the effects of both perceived recommendation quality ($\beta = -.39$, $p < .01$) and RA satisfaction ($\beta = -.08$, $p < .05$) on future use intentions. *Direct input from consumers* weakens the effect of perceived recommendation quality on RA satisfaction ($\beta = -.10$, $p < .05$) and future use intentions ($\beta = -.19$, $p < .01$). Finally, *personalized recommendations* enhance the effect of RA satisfaction on future use intention ($\beta = .24$, $p < .01$).

General discussion

Given advances in recommendation technology, we conducted a comparative assessment of different RA types in terms of the underlying algorithm, recommendation presentation, and data source. The proposed meta-analytic framework features the direct effect of perceived recommendation quality on consumers' future RA use intentions, with decision-making and RA satisfaction as mediators. We explore how RAs perform in leveraging the effects of perceived recommendation quality and the two mediators. Our findings have clear implications for RA theory and provide insights for managers in evaluating and selecting the best RAs.

Which RA types leverage the effects of perceived recommendation quality?

The results of the effect size integration suggest that perceived recommendation quality is positively related to decision-making and RA satisfaction, as well as future use intentions. We assessed the performance of RAs in terms of leveraging the effects of perceived recommendation quality on these mediators and outcomes. Although prior literature indicates performance differences of RA types (Table 1), the specific moderating effects are unclear. Thus, we followed the empirics-first approach to explore which RA types work best in supporting consumers. We classify RAs in terms of the underlying algorithm, recommendation presentation, and data source to assess performance differences.

First, regarding the *RA algorithm, collaborative filtering* performs best, as it enhances the effects of perceived recommendation quality on RA satisfaction and future use intentions. However, we do not find a difference for decision-making satisfaction. Scholars should explore whether certain consumers find collaborative filtering more diagnostic, because this algorithm cannot really provide recommendations for new consumers. *Content-based filtering* RAs perform worse than other algorithms, such that they weaken the effects of perceived recommendation quality on both satisfaction mediators, possibly due to the RA's inability to recommend new products beyond consumers' own preferences (Ricci, Rokach and Shapira 2022). We observe no differences in performance for *interactive RAs*, and we thus call for more research to assess when these RAs might be preferable. Finally, *self-serving recommendations* enhance the effects of perceived recommendation quality on decision-making satisfaction and future use intentions. Although these recom-

[3] We test potential nonlinear effects by coding an additional comprehensive variable (1 = medium vs. 0 = low/high comprehensiveness). This variable was nonsignificant for all five relationships ($p > .05$).

Table 4
Results of subgroup analysis for relationships in framework.

		Effects of RQ on Mediators and Future Use Intention									Effects of Mediators on Future Use Intention					
		RQ → Decision-making satisfaction			RQ → RA satisfaction			RQ → Future use intention			RA satisfaction → Future use intention			Decision-making satisfaction → Future use intention		
Moderator	Level	k	rwc	p	k	rwc	p	k	rwc	p	k	rwc	p	k	rwc	p
RA Algorithm																
Collaborative filtering	Yes	11	.43*	.14	14	.65*	.00	18	.64*	.00	21	.60*	.30	12	.54*	.00
	No	10	.39*		29	.43*		21	.41*		42	.58*		11	.31*	
Content-based filtering	Yes	15	.37*	.00	27	.44*	.00	27	.44*	.00	50	.61*	.00	16	.35*	.00
	No	6	.48*		16	.55*		12	.58*		13	.54*		7	.54*	
Interactive RA	Yes	—	—	—	7	.47*	.14	4 [b]	.34*	.00	18	.65*	.00	1 [a,b]	−.11	.00
	No	—	—		36	.50*		35	.54*		45	.56*		22	.45*	
Self-serving recommendation	Yes	2 [b]	.83*	.00	3 [b]	.60*	.00	3 [b]	.63*	.00	2 [b]	.53*	.14	1 [a,b]	.88*	.00
	No	19	.37*		40	.48*		36	.47*		61	.59*		22	.39*	
RA Presentation																
Solicited recommendation	Yes	3 [b]	.54*	.00	12	.40*	.00	8	.35*	.00	34	.60*	.08	4 [b]	.13*	.00
	No	18	.39*		31	.54*		31	.57*		29	.58*		19	.48*	
Comprehensiveness of results	Many	10	.40*	.57	22	.48*	.41	23	.50*	.32	35	.55*	.00	10	.40*	.16
	Few	11	.42*		21	.49*		16	.48*		28	.64*		13	.44*	
Website-embedded RA	Yes	7	.30*	.00	15	.47*	.41	12	.49*	1.00	33	.57*	.01	10	.40*	.17
	No	14	.45*		28	.49*		27	.49*		30	.61*		13	.44*	
RA Data																
Location-based RA	Yes	2	.69*	.00	4 [b]	.38*	.00	3 [b]	.60*	.00	4 [b]	.49*	.00	2 [b]	.74*	.00
	No	19	.37*		39	.50*		36	.48*		59	.60*		21	.38*	
Social network-based RA	Yes	1[a,b]	.46*	.42	4 [b]	.60*	.00	2 [b]	.24*	.00	4 [b]	.50*	.00	1 [a,b]	.52*	.13
	No	20	.41*		39	.48*		37	.50*		59	.60*		22	.42*	
Direct input from consumers	Yes	6	.49*	.00	16	.38*	.00	18	.43*	.00	39	.60*	.43	7	.32*	.00
	No	15	.38*		27	.55*		21	.55*		24	.58*		16	.46*	
Personalized recommendation	Yes	17	.38*	.00	34	.50*	.01	33	.49*	.71	57	.60*	.00	19	.41*	.14
	No	4 [b]	.49*		9	.44*		6	.50*		6	.52*		4 [b]	.45*	

* $p < .05$ (two-tailed). a. We report these effect sizes for sake of completeness; moderators were included in meta-regression when at least two observations were reported. b. Meta-analyses differ in the minimum number of effect sizes they use for moderator tests ranging from 2 to 5; thus, we highlight moderators with less than 5 effect sizes.

Table 5
Results of meta-regression for relationships in framework.

	Effects of RQ on Mediators and Future Use Intention						Effects of Mediators on Future Use Intention			
	RQ → Decision-making satisfaction		RQ → RA satisfaction		RQ → Future use intention		RA satisfaction → Future use intention		Decision-making satisfaction → Future use intention	
Moderator	B	p	B	p	B	p	B	p	B	p
Constant	.44*	.00	.61*	.00	.76*	.00	.30*	.00	.76*	.00
RA Algorithm										
Collaborative filtering (1=yes, 0=no)	—		.23*	.00	.07*	.02	—		.23*	.00
Content-based filtering (1=yes, 0=no)	−.17*	.02	−.10*	.02	−.06	.10	−.01	.84	−.24*	.00
Interactive RA (1=yes, 0=no)	—		—		—		−.01	.76	—	
Self-serving recommendation (1=yes, 0=no)	.25*	.05	−.10	.07	.11*	.01	—		—	
RA Presentation										
Solicited recommendation (1=yes, 0=no)	.34*	.00	−.02	.59	−.26*	.00	—		−.23*	.00
Comprehensiveness of results (1=many, 0=few)	—		—		—		−.07*	.01	—	
Website-embedded RA (1=yes, 0=no)	.08	.26	—		—		.02	.42	—	
RA Data										
Location-based RA (1=yes, 0=no)	.04	.61	−.10*	.04	.23*	.00	−.03	.53	.02	.79
Social network-based RA (1=yes, 0=no)	—		.24*	.00	−.39*	.00	−.08*	.04	—	
Direct input from consumers (1=yes, 0=no)	.03	.65	−.10*	.04	−.19*	.00	—		.06	.30
Personalized recommendation (1=yes, 0=no)	.07	.29	.05	.31	—		.24*	.00	—	
Controls[a]	Included		Included		Included		Included		Included	
R^2	.78		.45		.56		.42		.82	
k	21		43		39		63		23	
Variance inflation factor	5.69		5.02		3.06		3.80		4.40	

* $p < .05$ (two-tailed). The table displays unstandardized regression coefficients. A positive (negative) regression coefficient indicates a stronger (weaker) positive relationship between perceived recommendation quality and intention to use when the moderator is high than when it is low. For example, the positive coefficient of collaborative filtering on RQ-RA satisfaction relationship indicates that the positive relationship between perceived recommendation quality and RA satisfaction is stronger for RAs using collaborative filtering than RAs not using this algorithm. a. The detailed results with controls are shown in Web Appendix E. RQ = recommendation quality.

mendations are biased and often favor the firm, consumers still appreciate them as a decision-making heuristic. Scholars should explore the potential negative effects of these algorithms on the firm's brand image, as well as whether and when consumers find this practice unethical; issues related to privacy, ethics, and fairness of RAs seem likely to gain importance (Wirtz et al. 2022). Although currently, little oversight limits how firms can engage in nudging consumers to make decisions without their knowledge, this situation is likely to change in the future.

Second, for *RA presentation*, RAs that rely on *solicited recommendations* enhance the effect of perceived recommendation quality on decision-making satisfaction. Solicited (vs. unsolicited) recommendations seem more helpful to consumers. However, because consumers attribute the perceived recommendation quality to their own effort, the effect of perceived recommendation quality on future use intentions is weaker. Scholars should further study consumers' effort attribution. We find no differences for *comprehensiveness of results*; it seems consumers have no preferences between limited recommended options and longer recommendation lists. Nor can we identify nonlinear effects. Thus, we hope continued studies examine possible nonlinear effects of comprehensiveness of results and the optimal number of recommendations explicitly. The data indicate no differences for *website-embedded RAs*. We surmise that, because firms often use RAs on both websites and mobile apps, the differences between these presentation types have become less apparent.

Third, regarding *RA data, location-based RAs* enhance the effect of perceived recommendation quality on future use intentions but weaken its effect on RA satisfaction. Such recommendations are explicit, and the use of location information is generally appreciated, but this single piece of information may not generate the best options for consumers. *Social network–based RAs* also display mixed results: They enhance the effect of perceived recommendation quality on RA satisfaction but weaken its effect on future use intentions. Even if such RAs can provide socially acceptable recommendations, privacy concerns discourage consumers from relying on them. *Direct input from consumers* weakens the effects of perceived recommendation quality on RA satisfaction and future use intentions; consumers appear to perceive the RA's effort as lower than their own during the search process. We observe no differences for *personalized recommendations*, so more research is needed to test different types of personalization and which consumers appreciate them.

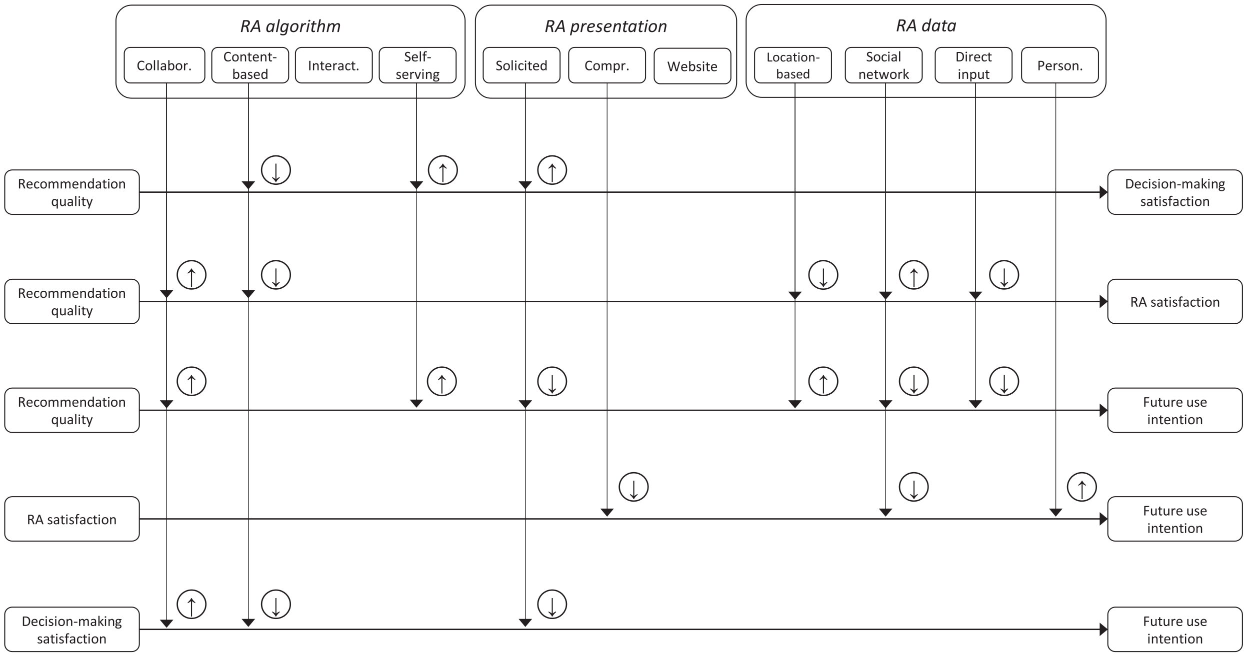

Fig. 3. Summary of RA performance assessment. Note: The figure displays only moderating effects that are significant in meta-regression; RAs without an arrow display an average performance as they neither enhance nor weaken any relationship. ↑= Enhancing effect; ↓= weakening effect.

Which RA types leverage the effects of mediators?

We also assess the performance of RA types in terms of leveraging the effects of mediators (i.e., decision-making and RA satisfaction) on future use intentions. Both mediators have positive effects; perhaps even more important, we observe several performance differences across the examined RA types.

First, among *RA algorithms, collaborative filtering* enhances the effect of decision-making satisfaction on future use intentions, whereas *content-based filtering* weakens it. The former outperforms other algorithms in translating satisfaction mediators into future use intentions, whereas the latter performs poorly, due to its limited ability to suggest novel products beyond consumers' known preferences. We observe no differences for *interactive RAs* and *self-serving recommendations*, and we recommend more research into their effects. For example, depending on their personal traits (e.g., technology readiness), some consumer may enjoy using interactive RAs and perceive them as fun, while others may regard their use as difficult.

Second, for *RA presentation*, we find that RAs that rely on *solicited recommendations* and *comprehensive results* presentations perform poorly and weaken the effect of decision-making and RA satisfaction on future use intentions. Scholars should explore whether these effects hold for all consumers (e.g., internet novices). A comprehensive results presentation also appears more likely to influence the effects of mediators on outcomes than the effects of perceived recommendation quality on mediators, though further exploration is needed to confirm this observation. We find no differences for *website-embedded RAs*, pointing again to the blurring line between website and nonwebsite platforms. Scholars might define the contexts in which website-embedded RAs perform better than those on other platforms.

Third, for *RA data*, we clarify that *personalized recommendations* enhance the effect of RA satisfaction on future use intentions; more research is needed to improve the predictive ability of RAs that leverage nonpersonalized data. *Social network–based RA* weakens the effect of RA satisfaction on future use intentions, given the information sensitivity of these data and consumers' privacy concerns. Scholars should explore whether some social networks (e.g., business networks like LinkedIn) provide more useful data, as well as try to identify consumers or products for which social network–based RAs might perform better. We uncover no differences for *location-based RAs* and *direct input from consumers*. Researchers should examine when consumers appreciate the use of location data and when they benefit more from providing direct input to RAs.

Managerial implications

Our study has several implications for practice. Fig. 3 summarizes the key findings, and Table 6 summarizes the implications. First, our study has clear implications for managers for selecting the best RA (Fig. 3). In terms of the number of enhancing/weakening effects, this meta-analysis shows that collaborative-filtering RAs perform best. The second-best options are RAs that provide self-serving and personalized recommendations. Interactive, website-embedded, and location-based RAs rank third. If RAs offer solicited recom-

Table 6
Managerial implications and research agenda on performance of RA types.

RA type	Managerial implications	Key illustrative recommendations
RA Algorithm		
Collaborative filtering	Firms should consider using RAs based on collaborative filtering, as this algorithm performs best among examined RAs.	Because we did not find a difference for decision-making satisfaction, scholars should assess whether it works equally well for different consumer groups (e.g., new consumers).
Content-based filtering	Managers must be careful when using RAs based on content-based filtering. This RA showed the worst performance.	These RAs lack the ability to recommend new products beyond consumers' own preferences; further research should determine whether it performs better in certain product categories.
Interactive RA	Firms can use interactive RAs that rely on AI; this new technology shows decent performance, though it does not (yet) outperform other algorithms.	Examine when and for which consumers the dynamic interaction with these RAs distracts consumers (e.g., consider consumer traits like computer playfulness or technology readiness).
Self-serving recommendation	Managers can use self-serving recommendations (sponsored search) but must consider potential negative side effects on brand image.	Explore when using these algorithms damages the firm's brand image, and should study privacy, ethics, and fairness of RAs as they gain importance in the future (Wirtz et al. 2022).
Multistakeholder RA	Not tested	Explore multistakeholder RAs that optimize consumer choices for multiple stakeholders, such as providers (e.g., hotels), system owners (e.g., Airbnb), and travel industry consumers.
Browsing RA	Not tested	Assess performance of RA algorithms that support browsing and search behavior without any intention of selling an item.
Consumer feedback	Not tested	Test performance of algorithms that use either explicit (e.g., ratings of purchased cars) or implicit consumer feedback.
RA Presentation		
Solicited recommendation	Introduce more RAs that provide unsolicited recommendations (e.g., during online checkout); they largely perform better than solicited recommendations.	Examine when, and which, consumers appreciate these RAs (e.g., internet novices, elderly consumers).
Comprehensiveness of results	Managers should be careful when using RAs that provide comprehensive results lists. RAs that provide short lists are preferable.	We observed few differences for this moderator; scholars should explore why this moderator is more likely to influence effects of mediators than perceived recommendation quality.
Website-embedded RA	RAs should be offered via not only websites but also different platforms (e.g., mobile/app-based RAs). Website-embedded RAs show average performance.	We observed no differences for this RA; scholars should explore which products are better recommended via websites and which via other devices.
Recommendation sequence	Not tested	Assess RAs that present sequences of items, one-by-one (e.g., next point of interest to visit) or as a whole (e.g., music compilation).
Recommendation of bundles	Not tested	Test RAs that present bundles including a group of products that fit well together (e.g., various attractions, destinations, and hotels).
Group recommender system	Not tested	Explore group recommender systems that present recommendations to groups of consumers (e.g., television programs).
RA Data		
Location-based RA	Managers should select RAs that use consumers' location data (e.g., GPS) on websites or mobile devices to improve certain outcomes.	Location-based RAs display some effects; however, there may be circumstances when consumers have more privacy concerns (e.g., health products).
Social network–based RA	Managers should be careful when employing RAs that use information from social networks. Privacy concerns discourage consumers from using these RAs.	Explore whether some social networks (e.g., LinkedIn) provide more useful data; explore consumers' privacy concerns for different data.
Direct input from consumers	These RAs perform worse than other RAs. Firms should consider collating consumers' data via means other than web interfaces and online forms.	We observed some differences for direct input from consumers; scholars should explore which kinds of data input from consumers lead to better RA evaluations.
Personalized recommendation	RAs providing unique recommendations based on personal consumer data should be preferred to RAs using nonpersonal data.	RAs that use personal data performed well; scholars should assess when to use nonpersonalized data and which data.
Natural language processing	Not tested	Assess the performance of RAs that use data generated through natural language processing from textual descriptions.
Session-based RA	Not tested	Examine session-based RAs that rely on data collected during the active session for consumers without a profile (e.g., guests of the system).
Further contextual data	Not tested	Test RAs that consider further contextual data such temporal, weather, and consumer mood data.

mendations, comprehensive results presentation, social media data, and direct input from consumers, they rank fourth; thus, managers should consider them more cautiously than the other types. Finally, RAs using content-based filtering algorithm rank last; they weaken many effects in our framework.

Second, our study suggests how to improve specific outcomes important to retailers (Fig. 3). Specifically, managers aiming to improve decision-making satisfaction should prioritize RAs that rely on self-serving and solicited recommendations but avoid those relying on content-based filtering. Managers who are more concerned about RA satisfaction should use collaborative filtering and social network–based RAs but avoid using those that rely on content-based filtering, location data, or direct input from consumers. If managers want to improve future use intentions, they should employ collaborative filtering, self-serving recommendations, and location-based RAs. They should avoid RAs that rely on solicited recommendations, social media data, and direct input from consumers. To improve the translation of RA satisfaction into future use intentions, managers should adopt personalized-recommendation RAs but avoid RAs that provide comprehensive results and rely on social media data. Finally, managers can enhance the translation of decision-making satisfaction into future use intentions by using collaborative-filtering RAs and avoiding those that rely on content-based filtering or solicited recommendations. Table 6 summarizes these implications.

Limitations and research agenda

Drawing from our meta-analysis, we propose a research agenda for studying the performance of RA types in Table 6. These recommendations take into account our empirical findings, along with limitations regarding the meta-analytic method and data availability.

First, we assessed the performance of RAs using several RA algorithms. We suggest exploring the performance of these algorithms for different, specific types of consumers and products. Although we assessed key RA algorithms (e.g., collaborative filtering), we could not include some of them, due to data availability constraints and their novelty. Continued research is needed into (1) multistakeholder RAs, in which providers (e.g., hotels) and system owners (e.g., Airbnb) jointly try to influence consumer choices; (2) RAs that support browsing and search behavior without any intention of making sales; and (3) RAs that use explicit or implicit consumer feedback (e.g., ratings of purchased cars) and other user-generated content to improve the algorithm.

Second, by building on our assessment of RA performance, according to four types of RA presentation, researchers might identify more types. Some previous studies mention other types, but they have not assessed the interplay with perceived recommendation quality. Continued studies could examine (1) the recommendation sequence—instead of generating a single recommendation, the RA suggests a sequence of items, one by one (e.g., next point of interest to visit) or as a whole (e.g., compilation of musical tracks); (2) recommendation bundles, including groups of related products (e.g., attractions, destinations, hotels); or (3) group recommender systems that present recommendations to a group of consumers rather than to an individual (e.g., television programs).

Third, we assessed the performance of RAs relying on various RA data, but other sources of knowledge could be informative as well, such as (1) data generated through natural language processing that extracts meaningful information from textual descriptions; (2) data generated by session-based RAs, in which data collected during an active session serve to make predictions for consumers' short-term needs, which could be useful for consumers without a registered profile (e.g., system guests, consumers who prefer anonymity); and (3) other contextual data such as temporal, weather, or consumer mood data. We hope that scholars find the proposed research agenda and our ideas in this exciting and important field inspiring for their continued studies.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jretai.2023.08.001.

References

Abumalloh, R.A., Ibrahim O. and Nilashi M. (2020), "Loyalty of Young Female Arabic Customers towards Recommendation Agents," *Technology in Society*, 61, 101253.

Aggarwal, P. and Vaidyanathan R. (2005), "Perceived Effectiveness of Recommendation Agent Routines: Search vs. Experience Goods," *International Journal of Internet Marketing and Advertising*, 2 (1), 38–55.

Anastasiya, Z. (2021), "Why Implement AI-Based Recommendation System," *InData Labs*. https://indatalabs.com/blog/ai-based-recommender-system. (accessed June 18, 2022).

Bendapudi, N. and Leone R.P. (2003), "Psychological Implications of Customer Participation in Co-Production," *Journal of Marketing*, 67 (1), 14–28.

Bhattacharyya, A., Jha S., Guha A. and Biswas A. (2022), "Should Firms Display the Sale Price Using Larger Font?," *Journal of Retailing*. doi.org/10.1016/j.jretai.2022.06.005.

Blut, M., Wang C., Wünderlich N.V. and Brock C. (2021), "Understanding Anthropomorphism in Service Provision: A Meta-Analysis of Physical Robots, Chatbots, and Other AI," *Journal of the Academy of Marketing Science*, 49 (4), 632–58.

Borenstein, M., Hedges L.V., Higgins J.P. and Rothstein H.R. (2009). *Introduction to Meta-Analysis*. John Wiley & Sons.

Borenstein, M., Hedges L.V., Higgins J.P. and Rothstein H.R. (2021). *Introduction to Meta-Analysis* (2nd ed). John Wiley & Sons.

Brown, S.P. and Lam S.K. (2008), "A Meta-Analysis of Relationships Linking Employee Satisfaction to Customer Responses," *Journal of retailing*, 84 (3), 243–55.

Castells, P., Hurley N. and Vargas S. (2022). "Novelty and Diversity in Recommender Systems". In *Recommender Systems Handbook* (pp. 603–646). New York, NY: Springer.

Chen, Y., Lu Y., Wang B. and Pan Z. (2019), "How do Product Recommendations Affect Impulse Buying?," *Information and Management*, 56 (2), 236–48.

Choi, J., Lee H.J. and Kim H.W. (2017), "Examining the Effects of Personalized App Recommender Systems on Purchase Intention: A Self and Social-Interaction Perspective," *Journal of Electronic Commerce Research*, 18 (1), 73–102.

Danckwerts, S., Meißner L. and Krampe C. (2020). "Hi, Can You Recommend a Movie?" Investigating Recommendation Chatbots in Media Streaming Services". *ECIS 2020-28th European Conference on Information Systems*.

Datta, H., van Heerde H.J., Dekimpe M.G. and Steenkamp J.B.E. (2022), "Cross-National Differences in Market Response: Line-Length, Price, and Distribution Elasticities in 14 Indo-Pacific Rim Economies," *Journal of Marketing Research*, 59 (2), 251–70.

DeLone, W.H. and McLean E.R. (1992), "Information Systems Success: The Quest for the Dependent Variable," *Information Systems Research*, 3 (1), 60–95.

Divyaa, L.R. and Pervin N. (2019), "Towards Generating Scalable Personalized Recommendations: Integrating Social Trust, Social Bias, and Geo-Spatial Clustering," *Decision Support Systems*, 122, 113066.

Ebrahimi, L., Mirabi V.R., Ranjbar M.H. and Pour E.H. (2019), "A Customer Loyalty Model for e-Commerce Recommendation Systems," *Journal of Information and Knowledge Management*, 18 (03), 1950036.

Filieri, R. (2015), "What Makes Online Reviews Helpful? A Diagnosticity-Adoption Framework to Explain Informational and Normative Influences in e-WOM," *Journal of Business Research*, 68 (6), 1261–70.

Fishbein, M. and Ajzen I. (1975). *Attitude, Intention and Behavior: An Introduction to Theory and Research*. Reading: Addison-Wesley.

Frank, B., Torrico B.H., Enkawa T. and Schvaneveldt S.J. (2014), "Affect Versus Cognition in the Chain From Perceived Quality to Customer Loyalty: The Roles of Product Beliefs and Experience," *Journal of Retailing*, 90 (4), 567–86.

Fytraki, A. (2018). *Behavioral Effects in Consumer Evaluations of Recommendation Systems*. Erasmus University Rotterdam Dissertation.

Ghasemaghaei, M. (2016). *Online Product Recommendation Agents Design: The Role of Cognitive Age and Agent Comprehensiveness*. McMaster University Doctoral Dissertation.

Ghiassaleh, A., Kocher B. and Czellar S. (2020), "Best Seller!? Unintended Negative Consequences of Popularity Signs on Consumer Choice Behavior," *International Journal of Research in Marketing*, 37 (4), 805–20.

Golder, P.N., Dekimpe M.G., An J.T., van Heerde H.J., Kim D.S. and Alba J.W. (2022), "Learning from Data: An Empirics-First Approach to Relevant Knowledge Generation," *Journal of Marketing*. doi:10.1177/00222429221129200.

Goodman, J.K., Broniarczyk S.M., Griffin J.G. and McAlister L. (2013), "Help or Hinder? When Recommendation Signage Expands Consideration Sets and Heightens Decision Difficulty," *Journal of Consumer Psychology*, 23 (2), 165–74.

Grewal, D., Puccinelli N. and Monroe K.B. (2018), "Meta-Analysis: Integrating Accumulated Knowledge," *Journal of the Academy of Marketing Science*, 46 (1), 9–30.

Herlocker, J.L., Konstan J.A., Terveen L.G. and Riedl J.T. (2004), "Evaluating Collaborative Filtering Recommender Systems," *ACM Transactions on Information Systems*, 22 (1), 5–53.

Hess, T.J., Fuller M. and Campbell D.E. (2009), "Designing Interfaces with Social Presence: Using Vividness and Extraversion to Create Social Recommendation Agents," *Journal of the Association for Information Systems*, 10 (12), 889–919.

Hong, W., Chan F.K., Thong J.Y., Chasalow L.C. and Dhillon G. (2014), "A Framework and Guidelines for Context-Specific Theorizing in Information Systems Research," *Information Systems Research*, 25 (1), 111–36.

Hostler, R.E., Yoon V.Y. and Guimaraes T. (2012), "Recommendation Agent Impact on Consumer Online Shopping: The Movie Magic Case Study," *Expert Systems with Applications*, 39 (3), 2989–99.

Huang, L.T. (2013), "Investigating the Role of Flow Experiences in Users' Reuse Intentions Toward Recommendation Agents: The Moderator or Product Knowledge," *IADIS International Journal on WWW/Internet*, 11 (3), 61–75.

Huang, L.T. (2016), "Exploring Utilitarian and Hedonic Antecedents for Adopting Information from a Recommendation Agent and Unplanned Purchase Behaviour," *New Review of Hypermedia and Multimedia*, 22 (1-2), 139–65.

Huang, J. and Zhou L. (2019), "The Dual Roles of Web Personalization on Consumer Decision Quality in Online Shopping," *Internet Research*, 29 (6), 1280–300.

Hung, Y.H., Hu P.C. and Lee W.T. (2013). "Improving the Design and Adoption of Travel Websites". *5th IASDR International Conference*.

Hunold, M., Kesler R. and Laitenberger U. (2020), "Rankings of Online Travel Agents, Channel Pricing, and Consumer Protection," *Marketing Science*, 39 (1), 92–116.

Hunter, J.E. and Schmidt F.L. (2004). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Sage.

Jannach, D., Lerche L. and Jugovac M. (2015), "Item Familiarity as a Possible Confounding Factor in User-Centric Recommender Systems Evaluation," *I-com*, 14 (1), 29–39.

Jiang, Z. and Benbasat I. (2004), "Virtual Product Experience: Effects of Visual and Functional Control of Products on Perceived Diagnosticity and Flow in Electronic Shopping," *Journal of Management Information Systems*, 21 (3), 111–47.

Khare, A., Labrecque L.I. and Asare A.K. (2011), "The Assimilative and Contrastive Effects of Word-of-Mouth Volume," *Journal of Retailing*, 87 (1), 111–26.

Köcher, S., Jugovac M., Jannach D. and Holzmüller H.H. (2019), "New Hidden Persuaders: An Investigation of Attribute-Level Anchoring Effects of Product Recommendations," *Journal of Retailing*, 95 (1), 24–41.

Komiak, S.Y. and Benbasat I. (2006), "The Effects of Personalization and Familiarity on Trust and Adoption of Recommendation Agents," *MIS Quarterly*, 941–60.

Ku, Y.C. and Tai Y.M. (2013). "What Happens When Recommendation System Meets Reputation System? The Impact of Recommendation Information on Purchase Intention". In *46th Hawaii International Conference on System Sciences* (pp. 1376–1383).

Kwon, Y., Park J. and Son J.Y. (2020), "Accurately or Accidentally? Recommendation Agent and Search Experience in Over-the-Top (OTT) Services," *Internet Research*, 31 (2), 562–86.

Lee, Y.E. and Benbasat I. (2011), "The Influence of Trade-off difficulty Caused by Preference Elicitation Methods on User Acceptance of Recommendation Agents Across Loss and Gain Conditions," *Information Systems Research*, 22 (4), 867–84.

Lee, S. and Choi J. (2017), "Enhancing User Experience with Conversational Agent for Movie Recommendation," *International Journal of Human–Computer Studies*, 103, 95–105.

Lee, G. and Lee W.J. (2009), "Psychological Reactance to Online Recommendation Services," *Information and Management*, 46 (8), 448–452.

Liao, S.H., Widowati R. and Chang H.Y. (2021), "A Data Mining Approach for Developing Online Streaming Recommendations," *Applied Artificial Intelligence*, 35 (15), 2204–27.

Lipsey, M.W. and Wilson D.B. (2001). *Practical Meta-Analysis*. Thousand Oaks: Sage Publications.

Maity, M., Dass M. and Malhotra N.K. (2014), "The Antecedents and Moderators of Offline Information Search: A Meta-Analysis," *Journal of Retailing*, 90 (2), 233–54.

Marchand, A. and Marx P. (2020), "Automated Product Recommendations with Preference-Based Explanations," *Journal of Retailing*, 96 (3), 328–43.

McKinsey. (2020). *The State of AI in 2020* https://www.mckinsey.com/business-functions/mckinsey-analytics/our-insights/global-survey-the-state-of-ai-in-2020 (accessed January 21, 2022).

McNee, S.M., Riedl J. and Konstan J.A. (2006). "Being Accurate is not Enough: How Accuracy Metrics have Hurt Recommender Systems". In *Human Factors in Computing Systems* (pp. 1097–1101).

Nilashi, M., Jannach D., bin Ibrahim O., Esfahani M.D. and Ahmadi H. (2016), "Recommendation Quality, Transparency, and Website Quality for Trust-Building in Recommendation Agents," *Electronic Commerce Research and Applications*, 19, 70–84.

Pillai, R. and Sivathanu B. (2020), "Adoption of AI-based Chatbots for Hospitality and Tourism," *International Journal of Contemporary Hospitality Management*, 32 (10), 3199–226.

Pu, P., Chen L. and Hu R. (2011). "A User-Centric Evaluation Framework for Recommender Systems". In *Proceedings of fifth ACM conference on Recommender Systems* (pp. 157–164).

Punj, G.N. and Moore R. (2007), "Smart Versus Knowledgeable Online Recommendation Agents," *Journal of Interactive Marketing*, 21 (4), 46–60.

Retailtouchpoints. (2021). *How many Products Does Amazon Carry?* https://retailtouchpoints.com/resources/how-many-products-does-amazon-carry (accessed June 22, 2021).

Ricci, F., Rokach L. and Shapira B. (2022). "Recommender Systems: Techniques, Applications, and Challenges". In *Recommender Systems Handbook* (pp. 1–35). New York, NY: Springer.

Rikitianskii, A., Harvey M. and Crestani F. (2014). "A Personalised Recommendation System for Context-Aware Suggestions". In *European Conference on Information Retrieval* (pp. 63–74). Cham: Springer.

Roschk, H., Loureiro S.M.C. and Breitsohl J. (2017), "Calibrating 30 Years of Experimental Research: A Meta-Analysis of the Atmospheric Effects of Music, Scent, and Color," *Journal of retailing*, 93 (2), 228–40.

Rosenthal, R. (1979), "The File Drawer Problem and Tolerance for Null Results," *Psychological Bulletin*, 86 (3), 638–41.

Roth, P.L., Le H., Oh I.S., Van Iddekinge C.H. and Bobko P. (2018), "Using Beta Coefficients to Impute Missing Correlations in Meta-Analysis Research: Reasons for Caution," *Journal of Applied Psychology*, 103 (6), 644–58.

Roudposhti, V.M., Nilashi M., Mardani A., Streimikiene D., Samad S. and Ibrahim O. (2018), "A New Model for Customer Purchase Intention in e-Commerce Recommendation Agents," *Journal of International Studies*, 11 (4), 237–53.

Schepers, J. and Wetzels M. (2007), "A Meta-Analysis of the Technology Acceptance Model: Investigating Subjective Norm and Moderation Effects," *Information & Management*, 44 (1), 90–103.

Schuetzler, R.M., Grimes G.M. and Giboney J.S. (2020), "The Impact of Chatbot Conversational Skill on Engagement and Perceived Humanness," *Journal of Management Information Systems*, 37 (3), 875–900.

Seiders, K., Voss G.B., Grewal D. and Godfrey A.L. (2005), "Do Satisfied Customers Buy More? Examining Moderating Influences in a Retailing Context," *Journal of Marketing*, 69 (4), 26–43.

Senecal, S. and Nantel J. (2004), "The Influence of Online Product Recommendations on Consumers' Online Choices," *Journal of Retailing*, 80 (2), 159–69.

Sethuraman, R., Gázquez-Abad J.C. and Martínez-López F.J. (2022), "The Effect of Retail Assortment Size on Perceptions, Choice, and Sales: Review and Research Directions," *Journal of Retailing*, 98 (1), 24–45.

Shi, S., Gong Y. and Gursoy D. (2021), "Antecedents of Trust and Adoption Intention Toward Artificially Intelligent Recommendation Systems in Travel Planning: A Heuristic–Systematic Model," *Journal of Travel Research*, 60 (8), 1714–34.

Tsekouras, D. and Li T. (2015). "The Dual Role of Perceived Effort in Personalized Recommendations". *Twenty-Third European Conference on Information Systems*.

Tsekouras, D., Li T. and Benbasat I. (2022), "Scratch My Back and I'll Scratch Yours: The Impact of User Effort and Recommendation Agent Effort on Perceived Recommendation Agent Quality," *Information and Management*, 59 (1), 103571.

Van Doorn, J., Mende M., Noble S.M., Hulland J., Ostrom A.L., Grewal D. and Petersen J.A. (2017), "Domo Arigato Mr. Roboto: Emergence of Automated Social Presence in Organizational Frontlines and Customers' Service Experiences," *Journal of Service Research*, 20 (1), 43–58.

Venkatesh, V., Morris M., Davis G.B. and Davis F.D. (2003), "User Acceptance of Information Technology: Toward a Unified View," *MIS Quarterly*, 27 (3), 425–78.

Verma, D., Dewani P.P., Behl A. and Dwivedi Y.K. (2023), "Understanding the Impact of eWOM Communication through the Lens of Information Adoption Model: A Meta-Analytic Structural Equation Modeling Perspective," *Computers in Human Behavior*, 107710.

Whang, C. and Im H. (2018), "Does Recommendation Matter for Trusting Beliefs and Trusting Intentions?," *International Journal of Retail & Distribution Management*, 46 (10), 944–58.

Whang, C. and Im H. (2021), ""I Like Your Suggestion!" The Role of Humanlikeness and Parasocial Relationship on the Website Versus Voice Shopper's Perception of Recommendations," *Psychology & Marketing*, 38, 581–95.

Wirtz, J., Patterson P.G., Kunz W.H., Gruber T., Lu V.N., Paluch S. and Martins A. (2018), "Brave New World: Service Robots in the Frontline," *Journal of Service Management*, 29 (5), 907–31.

Wirtz, J., Tarbit J.J., Hartley N. and W (2022), "Corporate Digital Responsibility: Dealing with Ethical, Privacy and Fairness Challenges of AI," *Journal of AI, Robotics & Workplace Automation*, 1 (4), 325–8.

Xiao, B. and Benbasat I. (2007), "E-commerce Product Recommendation Agents: Use, Characteristics, and Impact," *MIS Quarterly*, 31 (1), 137–209.

Xiao, B. and Benbasat I. (2018), "An Empirical Examination of the Influence of Biased Personalized Product Recommendations on Consumers' Decision Making Outcomes," *Decision Support Systems*, 110, 46–57.

Xu, J., Benbasat I. and Cenfetelli R.T. (2014), "The Nature and Consequences of Trade-Off Transparency in the Context of Recommendation Agents," *MIS Quarterly*, 38 (2), 379–406.

Yen, C. and Chiang M.C. (2021), "Trust Me, if You can: A Study on the Factors that Influence Consumers' Purchase Intention Triggered by Chatbots Based on Brain Image Evidence and Self-Reported Assessments," *Behaviour and Information Technology*, 40 (11), 1177–94.

Yoon, V.Y., Hostler R.E., Guo Z. and Guimaraes T. (2013), "Assessing the Moderating Effect of Consumer Product Knowledge and Online Shopping Experience on Using Recommendation Agents for Customer Loyalty," *Decision Support Systems*, 55 (4), 883–93.

Zhang, B. (2016). *Can Customization of Privacy Settings Promote Persuasiveness of Personalized Recommendation Agents?*. Dissertation. Pennsylvania State University.

Zhang, T., Agarwal R. and Lucas Jr H.C. (2011), "The Value of IT-Enabled Retailer Learning: Personalized Product Recommendations and Customer Store Loyalty in Electronic Markets," *MIS Quarterly*, 859–81.

Zhang, H., Zhao L. and Gupta S. (2018), "The Role of Online Product Recommendations on Customer Decision Making and Loyalty in Social Shopping Communities," *International Journal of Information Management*, 38 (1), 150–66.

Zhu, D., Chang Y.P., Luo J.J. and Li X. (2014), "Understanding the Adoption of Location-based Recommendation Agents among Active Users of Social Networking Sites," *Information Processing and Management*, 50 (5), 675–82.

Zhu, D., Sun H. and Chang Y. (2017), "How the Content of Location-based Advertisings Influences Consumers' Store Patronage Intention," *Journal of Consumer Marketing*, 34 (7), 603–11.