

2014

“Do you trust me?” – A Structured Evaluation of Trust and Social Recommendation Agents

Nicolai Walter

University of Muenster - ERCIS, nicolai.walter@uni-muenster.de

Follow this and additional works at: <http://aisel.aisnet.org/sighci2014>

Recommended Citation

Walter, Nicolai, “Do you trust me?” – A Structured Evaluation of Trust and Social Recommendation Agents” (2014). *SIGHCI 2014 Proceedings*. 21.

<http://aisel.aisnet.org/sighci2014/21>

This material is brought to you by the Special Interest Group on Human-Computer Interaction at AIS Electronic Library (AISeL). It has been accepted for inclusion in SIGHCI 2014 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

“Do you trust me?” – A Structured Evaluation of Trust and Social Recommendation Agents

Nicolai Walter

University of Muenster - ERCIS
nicolai.walter@uni-muenster.de

ABSTRACT

Recommender systems are considered as useful software that helps users in screening and evaluating products. The fact that users do not know how these systems make decisions leads to an information asymmetry. Thus, users need to trust if they want to take over systems' recommendations. Applying social interfaces has been suggested as helpful extensions of recommender systems to increase trust. These are called (Social) Recommendation Agents. While many articles and implementations can be found in the field of e-commerce, we believe that Recommendation Agents can be applied to other contexts, too. However, a structured evaluation of contexts and design dimensions for Recommendation Agents is lacking. In this study, first, we give an overview of design dimensions for Recommendation Agents. Second, we explore previous research on trust and Recommendation Agents by means of a structured literature review. Finally, based on the resulting overview, we highlight three major areas for future research.

Keywords

Social Recommendation Agents, Social Presence, Trust.

INTRODUCTION

A recommender system is a piece of software that provides helpful suggestions for users based on their preferences. For example, in e-commerce recommender systems have become popular because users may delegate the task of product screening and evaluation to them (Xiao and Benbasat 2007). While this process results in a presentation of recommendations, the underlying selection process itself remains hidden from users' perspective. This inherent information asymmetry shows that relationships between users and recommender systems can be described as agency relationships (Xiao and Benbasat 2007). From a user's perspective who does not know how recommendations evolve, trust in the system is needed in order to follow its suggestions.

When it comes to online trust research, Social Presence (SP), i.e., the feeling of human warmth, (Short, Williams, and Christie 1976), is considered to be an important influence factor which can be fostered by IT design decisions (Hess, Fuller, and Campbell 2009). In terms of

recommender systems, SP usually refers to a type of “animated embodiments [i.e., visual, often human-looking representations] that respond to users through verbal and nonverbal communication” (Chattaraman et al. 2012, p. 2055). This kind of virtual agent is called (*Social Recommendation Agent* (RA) (Hess et al. 2009). While recommender systems have been extensively studied (Xiao and Benbasat 2007), a structured evaluation of RAs has been neglected so far.

THEORETICAL BACKGROUND

Designing Social Presence

Studies have shown that users respond to computers as they were people (Nass and Moon 2000). This so called computer as social actor (CASA) paradigm states that users apply the same social rules to computers as in personal interactions. The general effect of social reactions to computers is even stronger when RAs are *embodied* (Nowak and Biocca 2003). This visual manifestation of RAs is also called “avatar” (Qiu and Benbasat 2009). Avatars have been shown to elicit feelings of SP (Nowak and Biocca 2003). Avatars do not necessarily have to be human-looking. The perception of SP and social responses will likely be stronger if representations of RAs are anthropomorphic (human-looking) (Nowak and Biocca 2003). There is a fast amount of design options. For example, RAs may be male or female, two-dimensional or three-dimensional, naturalistic or cartoonized, high detailed or degraded in levels of detail (Swinth and Blascovich 2002).

Besides embodiments, RAs are quasi-humans that entail both, technical aspects, and human characteristics. Thus, for designing RAs a set of distinct design dimensions should be considered (Nowak and Biocca 2003; Swinth and Blascovich 2002). Regarding technical aspects, RAs may use communication in order to interact with users. This potential to exert SP varies along two dimensions, namely vividness and interactivity (Fortin and Dholakia 2005; Steuer 1992). *Vividness* is defined as “the ability of a technology to produce a sensorially rich mediated environment” (Steuer 1992, p. 80). Thus, vividness refers to the degree to which a medium allows to convey social cues. This capability is determined by the number of different sensory channels (“breath”) and the resolution and quality of a particular channel (“depth”) (Steuer

1992). Studies of RAs may therefore vary across “breath” (e.g., text vs. audio output), or “depth” (e.g., automated text-to-speech [TTS] engines vs. recorded human voice). *Interactivity* is considered as “the degree to which users of a medium can influence the form and content of the media environment” (Steuer 1992, p. 80). This refers to the degree to which users can bring reciprocal effects to RAs such as turn taking and feedback (Skalski and Tamborini 2007). More interactivity likely leads to higher levels of perceived SP, as RAs “by virtue of their ability to interact with computer users, are proffered to instill a greater sense of SP than nonsocial agents” (Skalski and Tamborini 2007, p. 387). Interactivity is based on the perception of the party in terms of visible and audible in- and output (e.g. text field) (Clark and Brennan 1993).

In terms of human characteristics, RAs may be designed with *personalities* which may affect their verbal and nonverbal behavior (Andrews 2012). In order to raise the perception of SP, a high level of extraversion and agreeableness has been suggested to be most relevant (Hess et al. 2009). RAs may share their hobbies, personal histories or relationship information. Personality may also affect the other RAs design dimensions as, for example, “extraversion [...] can be recognized in both verbal and nonverbal cues” (Hess et al. 2009, p. 894). Thus, personality may influence RAs vividness (e.g., RAs with extraverted voices), interactivity (e.g., asking more questions), and nonverbal behavior (e.g., more gestures).

As facial expressions alone can provide more than 50% of the meaning of a message, *nonverbal cues* can be seen as distinct design dimension for RAs (Cowell and Stanney 2005). RAs that possess human-like physical properties are able to transmit nonverbal signals (Vogele and Bente 2010). Fully embodied RAs may use gestures such as open arms in order to welcome users. Besides movements with arms and legs, “in particular, the animated face may exhibit a number of nonverbal cues [...] which contribute to feelings of social interactions” (Qiu and Benbasat 2009, p. 150). There is a broad range of nonverbal cues, e.g., eye contact, mirroring, pointing, and expressive facial expressions such as lip movements and smiles (Vogele and Bente 2010; Qiu and Benbasat 2009).

Trust

Trust can be defined as “the willingness of a party [trustor] to be vulnerable to the actions of another party [trustee] based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party” (Mayer et al. 1995, p. 712). Online environments include environmental uncertainties that create risks (Pavlou 2003). Such risks could refer to the abuse of private data or penetration of one’s private sphere, or loss of money in online transactions. Dealing with RAs in an e-commerce transaction, risks may refer to “the user’s perceptions of uncertainty and potentially adverse consequences of buying a recommended product” (Xiao

and Benbasat 2007, p. 145). In such agency relationship with underlying information asymmetry users cannot determine whether RAs are capable of performing the tasks delegated to them (Xiao and Benbasat 2007).

The assessment of the trustworthiness of another party differs between initial *situations* and situations based on former contacts with the trustee, also called ongoing trust (McKnight, Choudhury, and Kacmar 2002). While in relationships trust is assessed on basis of experiences with the other party, in initial trust situations trust is based on the individual disposition of trust and the evaluation of the perceived cues provided in the initial situation (Mayer, Davis, and Schoorman 1995). The actual assessment of trust, i.e., forming trusting beliefs about a certain trustee, is conceived as a multidimensional construct (Mayer, Davis, and Schoorman 1995). The three trusting beliefs are ability, benevolence and integrity. RAs need to be capable of giving good recommendations (ability), act in users’ interest (benevolent), and be predictable (integrity). Despite the fact that RAs are usually embedded in websites or software applications, the *trust object*, i.e., the party to whom trusting beliefs refers to is not always clear; trust may either be assessed in RAs, websites, or hosting companies respectively.

RESEARCH METHODOLOGY

In order to learn about the state of the art of research on embodied RAs and trust, a structured literature review in Web of Science and ACM Digital Library was conducted. Selected articles were analyzed along the identified design dimensions as stated in the theoretical background. The search term included “trust” and synonyms used to describe RAs (e.g., “virtual agent”, “shopping advisor”, “avatar”, or “decision aid”). The search covered title, abstract and keywords in order to come up with a broad initial set of articles. In order to extract relevant studies, selection criteria were applied. First, studies had to be “empirical” in the sense that “the study has to involve actual use of a RA (prototype or operational, web-based or stand-alone) by human users [...] in either online or physical settings” (Xiao and Benbasat 2007, p. 140). Second, RAs needed to be embodied at some point of time and take some sort of a role like advisor, assistant, counselor, or guide. Third, the study had to include trust or trustworthiness as dependent variable.

While the search terms yielded a broad set of studies, a very high number of false hits had to be excluded during the screening process. The main reason was that “avatar” is used frequently for visual representation of other human actors in virtual worlds (e.g., Linden Lab’s Second Life). Moreover, “decision aid” also has a broader meaning (e.g. for technical devices in the healthcare sector). Finally, some studies referred to RAs with respect to algorithms that did not include any kind of embodiment or social aspects. The numbers of initial hits as of May 28th, 2014, were 158 (Web of Science) and 58 (ACM Digital Library). 11 studies were finally extracted.

RESULTS

Tables 1 and 2 show an overview of the results. Experimental manipulations are indicated by an asterisk. Hypothesized effects are ordered, e.g., “voice/text” means that voice is hypothesized to create more SP than text.

Author (Year)	Context	Designing SP	
		Embodiment	Vividness
Richards and Bransky (2014)	Real estate	Naturalistic 3D humanoid (female)	TTS
Lisetti et al. (2013)	Electronic health	*[n.s.] Naturalistic 3D humanoid (female)	Text and TTS
Chattaraman et al. (2012)	Electronic commerce	*[partly sig.] Naturalistic 3D humanoid (female)	Text and TTS
Pak et al. (2012)	Electronic Health	*[partly sig.] Naturalistic 3D humanoid (female)	Text
Al-Natour et al. (2011)	Electronic commerce	Naturalistic 2D humanoid (male, female)	Text, voice
Hess et al. (2009)	Real estate	Cartoonized 3D humanoid (male)	*[sig.] Text, TTS / Text
Qiu and Benbasat (2009)	Electronic commerce	*[sig.] Naturalistic 2D humanoid (male)	*[partly sig.] Human voice/ TTS/Text
Van Vugt et al. (2009)	Electronic health	*Cartoonized 2D humanoid (male, female): [n.s.] similarity; [n.s.] idealness	Text
Al-Natour et al. (2008)	Electronic commerce	Naturalistic 2D humanoid (male)	Text, voice
Cowell and Stanney (2005)	Private (photo sorting software)	Naturalistic 3D humanoid (male, female)	Text
Qiu and Benbasat (2005)	Electronic commerce	*[n.s.] Naturalistic 3D humanoid (male)	*[partly sig.] Text, TTS/ TTS/Text

Table 1. Overview of the results

Regarding the interactivity dimension, “low” refers to RAs that only sent information (one-way), “medium” characterizes RAs that asked questions or reacted to a

predefined set of questions, “high” was assigned to RAs that could be asked questions without any obvious restrictions. However, given the reported information in the studies, only an estimation of the quality of interactivity and nonverbal cues could be made.

Author (Year)	Designing SP (continued)		Trust	
	Inter-activity	Personality and nonverbal cues	Object	Situation
Richards and Bransky (2014)	Medium	Personality: *[partly sig.] Memory performance	RA	O
Lisetti et al. (2013)	Medium	Nonverbal cues: *[sig.] Facial expressions	RA	I
Chattaraman et al. (2012)	High	-	W	I
Pak et al. (2012)	Medium	-	RA	I
Al-Natour et al. (2011)	Low	Personality: *[n.s.] Personality similarity; *[sig.] Decision-making strategy similarity	RA	I
Hess et al. (2009)	Low	Personality: *[sig.] extraverted / introverted Nonverbal cues: *[sig.] Gestures	RA	I
Qiu and Benbasat (2009)	Medium	-	RA	I
Van Vugt et al. (2009)	High	-	RA	I
Al-Natour et al. (2008)	Low	Personality: *[sig.] Decision process similarity; *[n.s.] Decision outcome similarity	RA	I
Cowell and Stanney (2005)	Low	Nonverbal cues: *[sig.] Facial express; *[n.s.] Gestures	RA	I
Qiu and Benbasat (2005)	High	Nonverbal cues: Basic gestures such as waving	RA	I

Abbr.: Website (W), Initial trust (I), Ongoing trust (O)

Table 2. Continued overview of the results

DISCUSSION AND CONCLUSION

Implications for theory and practice

This article may support researchers in the field in better understanding and distinguishing RA design dimensions. The review helps to gain insight into the main advances in the field. Furthermore, researchers may take this as an overview in order to posit their research and connect their approaches to previous findings in the field. Finally, due to the mixed results and under-researched RA dimensions, this study highlights research gaps and makes suggestions for future research opportunities. Also practitioners like website designers can benefit from this structured evaluation of previous literature. While embodiments do not always lead to higher SP and trust, nonverbal cues seem to be promising design options that, in the best case, help, or, in the worst case, do neither improve nor harm. Thus, including sophisticated RAs with social abilities, nonverbal communication, and extroverted and agreeable personality seems to be a suitable option for many companies.

This study is not free of limitations. First, the presentation of the results did not explicitly consider the SP construct. Some studies did not measure SP because an influence on SP by including RAs is simply taken for granted (Chattaraman, Kwon, and Gilbert 2012). Second, we did not explicitly look at the role of participants. A detailed analysis of the effect of socio-demographic variables such as gender, age and cultural background would be interesting for future research. Finally, conference papers were not covered by our literature review.

Future Research

Besides addressing the limitations of this study by analyzing the role of participants and including conference papers, there are also many more promising contexts to study. As RAs provide both, recommender systems and social interfaces that may help in trust-building, promising contexts are such with high uncertainties and information need. In situation with product risks RAs may help in selecting suitable products and creating a positive attitude. Examples are flight and hotel booking, insurance products, or banking and investment consulting. Besides giving recommendations for comparing products, RAs may help users to gain deeper insight into a specific product or service. For instance, on websites of cloud computing providers, RAs may help to guide users through Service-Level-Agreements, security and privacy policies as well as data storage locations. In addition, context specific RA design requirements need to be studied in order to match contexts and RAs. For example, embodiment (e.g., clothing) and personality (e.g., use of language, level of discretion) may need to be designed very differently depending on hedonistic versus utilitarian contexts.

Many RA design dimensions show mixed results or no results. For example, despite the fact that *interactivity* is

an important antecedent of SP, not a single study varied systematically across different levels of interactivity. This may be due to the fact that it is challenging to implement RAs which are capable of answering questions. More precisely, questions to RAs could refer to both, the domain of expertise as well as social life. In terms of designing smooth answering behavior, various potential user questions need to be taken care of. This can be a time-consuming and costly endeavor. Thus, it would be very interesting to see how efforts of increasing interactivity paid off in terms of trust. Moreover, there are technical developments that allow for new design choices. For example, Microsoft's Kinect allows for motion-based interactivity. Other technologies such as Apple's Siri allow for an analysis of the human voice. Thus, future studies may experiment with both, voice input and reactions to users' gestures. Also there may be new contexts where such developments would be suitable. For example, welcome screens in buildings such as at university campuses where RAs could help visitors in finding directions.

Finally, future research for RAs and *trust* is needed. Regarding trust objects, most studies measured trust directly towards RAs. In theory, the relationship between embedded objects in websites, the websites, and the hosting company is not well explored. It would be interesting to see how trust differs across these measurement objects within a single study. With one exception, current studies only focus on initial trust. It would be important for researchers and companies to see if trust in RAs was maintained in the case of re-visits. Such a focus also offers new use cases. For example, RAs might save previously stated users' preferences and continually adapt their behavior on basis of user information. As privacy is at risk in such cases, the role of trust would also gain in importance.

REFERENCES

1. Al-Natour, S, I Benbasat, and R T Cenfetelli. 2008. "The Effects of Process and Outcome Similarity on Users' Evaluations of Decision Aids." *Decision Sciences*, 39, 2, 175–211.
2. Al-Natour, S, Izak Benbasat, and R. T. Cenfetelli. 2011. "The Adoption of Online Shopping Assistants: Perceived Similarity as an Antecedent to Evaluative Beliefs." *Journal of the Association for Information Systems*, 12, 5, 347–74.
3. Andrews, Pierre Y. 2012. "System Personality and Persuasion in Human-Computer Dialogue." *ACM Transactions on Interactive Intelligent Systems*, 2, 2, 1–27.
4. Chattaraman, Veena, Wi-Suk Kwon, and Juan E Gilbert. 2012. "Virtual Agents in Retail Web Sites: Benefits of Simulated Social Interaction for Older Users." *Computers in Human Behavior*, 28, 6, 2055–66.

5. Clark, H.H., and S.E. Brennan. 1993. *Grounding in Communication*. In: Baecker, R. M. (ed.), *Readings in Groupware and Computer-Supported Cooperative Work, Assisting Human-Human Collaboration*.
6. Cowell, A J, and K M Stanney. 2005. "Manipulation of Non-verbal Interaction Style and Demographic Embodiment to Increase Anthropomorphic Computer Character Credibility." *International Journal of Human-Computer Studies*, 62, 2, 281–306.
7. Fortin, David R, and Ruby Roy Dholakia. 2005. "Interactivity and Vividness Effects on Social Presence and Involvement with a Web-based Advertisement." *Journal of Business Research*, 58, 3, 387–96.
8. Goldberg, Lewis R. 1990. "An Alternative 'Description of Personality': The Big-Five Factor Structure." *Journal of Personality and Social Psychology*, 59, 6, 1216–1229.
9. Hess, Traci J, M Fuller, and D Campbell. 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.
10. Lisetti, Christine, Reza Amini, Ugan Yasavur, and Naphtali Rische. 2013. "I Can Help You Change! An Empathic Virtual Agent Delivers Behavior Change Health Interventions." *ACM Transactions on Management Information Systems*, 4, 4, 1–28.
11. Mayer, R C, J H Davis, and F D Schoorman. 1995. "An Integrated Model of Organizational Trust." *Academy of Management Review*, 20, 3, 709–34.
12. McKnight, D H, V Choudhury, and C Kacmar. 2002. "The Impact of Initial Consumer Trust on Intentions to Transact with a Web Site: a Trust Building Model." *Journal of Strategic Information Systems*, 11, 3-4, 297–323.
13. Nass, Clifford, and Y Moon. 2000. "Machines and Mindlessness: Social Responses to Computers." *Journal of Social Issues*, 56, 1, 81–103.
14. Nowak, K L, and F Biocca. 2003. "The Effect of the Agency and Anthropomorphism on Users' Sense of Telepresence, Copresence, and Social Presence in Virtual Environments." *Presence-Teleoperators and Virtual Environments*, 12, 5, 481–94.
15. Pak, R, N Fink, M Price, B Bass, and L Sturre. 2012. "Decision Support Aids with Anthropomorphic Characteristics Influence Trust and Performance in Younger and Older Adults." *Ergonomics*, 55, 9, 1059–72.
16. Pavlou, Paul A. 2003. "Consumer Acceptance of Electronic Commerce: Integrating Trust and Risk with the Technology Acceptance Model." *International Journal of Electronic Commerce*, 7, 3, 101–34.
17. Qiu, Lingyun, and I Benbasat. 2005. "Online Consumer Trust and Live Help Interfaces: The Effects of Text-to-speech Voice and Three-dimensional Avatars." *International Journal of Human-Computer Interaction*, 19, 1, 75–94.
18. Qiu, Lingyun, and Izak Benbasat. 2009. "Evaluating Anthropomorphic Product Recommendation Agents: A Social Relationship Perspective to Designing Information Systems." *Journal of Management Information Systems*, 25, 4, 145–82.
19. Richards, Deborah, and Karla Bransky. 2014. "ForgetMeNot: What and How Users Expect Intelligent Virtual Agents to Recall and Forget Personal Conversational Content." *International Journal of Human-Computer Studies*, 72, 5, 460–76.
20. Short, J., E. Williams, and B. Christie. 1976. *The Social Psychology of Telecommunications*. John Wiley & Sons, Ltd.
21. Skalski, Paul, and Ron Tamborini. 2007. "The Role of Social Presence in Interactive Agent-based Persuasion." *Media Psychology*, 10, 3, 385–413.
22. Steuer, Jonathan. 1992. "Defining Virtual Reality: Dimensions Determining Telepresence." *Journal of Communication*, 42, 4, 73–93.
23. Swinth, K, and Jim Blascovich. 2002. "Perceiving and Responding to Others: Human-human and Human-computer Social Interaction in Collaborative Virtual Environments." *The 5th Annual International Workshop on Presence by the International Society for Presence Research (ISPR)*.
24. Van Vugt, H C, E A Konijn, J F Hoorn, and J Veldhuis. 2009. "When Too Heavy Is Just Fine: Creating Trustworthy E-health Advisors." *International Journal of Human-Computer Studies*, 67, 7, 571–83.
25. Vogeley, K, and G Bente. 2010. "'Artificial Humans': Psychology and Neuroscience Perspectives on Embodiment and Nonverbal Communication." *Neural Networks*, 23, 8-9, 1077–90.
26. Xiao, B, and I Benbasat. 2007. "E-commerce Product Recommendation Agents: Use, Characteristics, and Impact." *MIS Quarterly*, 31, 1, 137–209.