

Are interface agents scapegoats? Attributions of responsibility in human–agent interaction

Alexander Serenko *

Faculty of Business Administration, Lakehead University, 955 Oliver Road, Thunder Bay, Ont., Canada P7B 5E1

Received 26 July 2005; received in revised form 25 July 2006; accepted 26 July 2006

Available online 12 September 2006

Abstract

This paper presents an investigation of the self-serving biases of interface agent users. An experiment that involved 202 MS Office users demonstrated that, in contrast to the self-serving hypothesis in attribution theory, people do not always attribute the successful outcomes of human–agent interaction to themselves and negative results to interface agents. At the same time, it was found that as the degree of autonomy of MS Office interface agents increases, users tend to assign more negative attributions to agents under the condition of failure and more positive attributions under the condition of success. Overall, this research attempts to understand the behavior of interface agent users and presents several conclusions that may be of interest to human–computer interaction researchers and software designers working on the incorporation of interface agents in end-user systems.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Interface agents; Human–agent interaction; Attribution theory; Self-serving bias

1. Introduction

Psychology researchers argue that it is natural for people to refer to their personal factors when they succeed and to blame other individuals or external circumstances when they fail. For example, when students successfully pass a test, they often mention their personal abilities, hard work, or good subject knowledge. However, when students fail, some of them tend to attribute this outcome to an unfair test or the quality of instruction. The psychology literature labels this phenomenon as self-serving bias – people tend to take personal credit when they succeed and deny their responsibility for failure. Some individuals also transfer their self-serving biases to modern technologies; they assign more responsibility to the negative effects of the use of machines (Postman, 1992). Currently, due to the diffusion of various software applications, people have developed a tendency to hold com-

puters morally responsible for errors (Friedman, 1995, 1997; Moon and Nass, 1998; Moon, 2003).

Computers are not only assistants, but also decision influencers or even decision-makers. For example, automated pilots, expert systems, and decision support applications have the potential to provide a user with the best available solution. At the same time, as computers become more autonomous (or independent), people may tend to increasingly attribute negative outcomes to machines and take credit for successful results (Kling, 1996).

Interface agents are an emerging technology that emphasizes the autonomy of software systems. An interface agent is a software entity which acts autonomously, monitors the environment, reacts to external stimuli, and communicates with users directly (Maes, 1994, 1995; Detlor, 2004). It exhibits strong visual and audio presence in the computer interface, and users are aware of the agent's existence (Serenko and Detlor, 2004; Serenko, 2006a). The goal of the incorporation of interface agents in end-user computer applications is to implement an indirect management approach instead of a less efficient direct manipulation method. Interface agents act as an intermediary between a

* Tel.: +1 807 343 8698; fax: +1 807 343 8443.

E-mail address: aserenko@lakeheadu.ca

person and various components of a computer system. They observe user actions taken in the direct manipulation interface, understand user needs, learn their behavior, create their profiles, and give advice (Lieberman, 2001; Lieberman and Selker, 2003). In contrast to reactive and predictable conventional direct manipulation interfaces, proactive interface agents work continuously and autonomously in the background by monitoring the external environment (Shneiderman and Maes, 1997). They act on behalf of the users by automatically invoking commands provided by software applications, by cooperating with other agents that constitute agent architecture, and by interacting with software processes. Interface agents initiate communication with users when they consider it necessary and appropriate. They may be employed in various forms, for instance, as virtual tutors (Lester et al., 1997; Johnson et al., 2000; Person et al., 2000), Web guides (Lieberman, 1995; Keeble and Macredie, 2000), or personal assistants (Maes and Kozierek, 1993; Lashkari et al., 1994). For the past decade, researchers have begun experimenting with the incorporation of animated, human, or cartoon-like interface agents in the graphical user interfaces of end-user systems (Ball et al., 1997). Currently, animated interface agents are incorporated in all Microsoft Office suites starting with Office 97 (Windows) and Office 98 (Macintosh).

Despite the potential advantages of the use of interface agents in MS Office, both academic literature and periodicals provide controversial evidence on the superiority of agent-mediated assistance (Serenko, 2007). For example, technologies that allow removing interface agents from MS Office were developed (Trott, 1998). Given negative user feedback on the constantly annoying animated Paper Clip, Microsoft limited the presence of animated help characters on the screen of Office XP (Magid, 2001). One of the reasons for negative user attitudes towards interface agents in MS Office may be the fact that interface agents are mostly associated with the negative consequences of application usage, whereas people take credit for successful work. In addition, since interface agents in MS Office seem to be autonomous, users may tend to over-attribute the negative outcomes of computer usage to agents. However, contemporary research offers no evidence to support this claim, and any documentation on studies of attributions of responsibility in human-agent interaction is missing. An understanding of the psychological factors leading to the acceptance or rejection of interface agents in MS Office is important given the magnitude of the issue. As such, Microsoft has already incorporated interface agents in the form of user assistants into MS Office and most people have become the involuntary users of this technology.

Therefore, this paper presents an empirical investigation of the self-serving biases of interface agent users to answer the following research question:

Do interface agent users attribute positive outcomes to their personal factors and negative outcomes to interface agents in human-agent interaction in MS Office?

2. The self-serving biases of interface agent users

Attribution theory has developed over time from several key works of Heider (1958), Jones and Davis (1965), and Kelly (1972). It explains how people make causal explanations about events and describes the behavioral outcomes of those explanations. While there are different perspectives within attribution theory (Peterson et al., 2002), most scholars would agree that the self-serving hypothesis (variously labeled as ego-defensive, ego-protective, or ego-biased attribution) states that individuals tend to assign differently the causes of their successes and failures in different situations; they engage in self-enhancing attributions under conditions of success and, on the contrary, in self-protective attributions under conditions of failure (Weiner et al., 1972; Miller and Ross, 1975). Causal attribution can be internal, associated with positive events ('I am responsible ...'), and external, associated with negative incidents ('Other people or situational factors are responsible ...') (Blackwood et al., 2003). The self-serving bias is a trait which is assumed to have the adaptive function of enhancing self-esteem.

Prior research confirms the presence of self-serving effects in various contexts, for example, in economic behavior (Farmer and Pecorino, 2002), decision-making (Roch et al., 2000), management (Gadhoun, 1999), corporate governance (Schwenk, 1993), organizational behavior (Lee and Tiedens, 2001), and information systems (Peterson et al., 2002). Moon and Nass (1998) and Moon (2003) proved the existence of self-serving biases in human-computer interaction.

Individuals tend to respond to computer technologies socially (Reeves and Nass, 1996); they use social rules in addressing computer behavior (Nass et al., 1994), apply human stereotypes to machines (Nass and Moon, 2000), and accept computers as teammates (Nass et al., 1996). Therefore, attribution theory may be employed to study the behavior of computer users (Moon and Nass, 1998). Since interface agents are software entities that interact with people directly and serve as their helpers, advisors, or assistants, it can be assumed that individuals apply social rules, behaviors, and expectations to interface agents to the same extent as they relate similar social principles to computers in general. Therefore, the area of social psychology offers a strong theoretical framework that may be successfully utilized to study human-agent interaction. Particularly, the self-serving hypothesis of attribution theory may be applied to investigate the behavior of interface agent users.

The investigation of the self-serving biases of interface agent users is important because there are both benefits and costs of external and internal attributions. On the one hand, self-serving behavior offers at least two advantages to attributors (Lee and Tiedens, 2001). First, individuals feel good about themselves by taking personal credits for success and protect their self-esteem by disclaiming responsibility for failure. Second, self-serving attributions

may serve public impressions. For example, when external causes are provided for misbehavior, people are less angry at the transgressor and express more positive attitudes towards the incident (Weiner et al., 1987).

On the other hand, there is evidence to suggest that engagement in self-serving behavior is not always beneficial because inadequate external attributions may have negative psychological effects on people in the long-run (Tennen and Affleck, 1990; Lee and Tiedens, 2001). Overall, as the power and control over the outcome of the third party increase, people tend to hold it more responsible for the final results. In the case of MS Office, if people tend to over-attribute their personal responsibility for unsuccessful computer tasks to interface agents, they may form negative attitudes towards using agents in general and eventually entirely reject this technology.

Presently, the types of user support provided by the MS Agent technology for Office applications are relatively limited. In general, two categories of assistance are offered: (1) support with a help menu and (2) the presentation of tips and real-time advice. These two kinds of assistance differ in their nature and in the way interface agents interact with users. In the former case, people explicitly initiate all actions by calling agents in a help menu. In the latter situation, agents pop up when the system believes it is necessary to offer user support.

As such, based on the self-serving hypothesis, it can be assumed that interface agent users act in a self-serving manner – they attribute every success in human–agent interaction to their internal factors, for example, to their computer skills, agent manipulation or field expertise, and blame an agent for every failure, for example, when the agent does not offer useful information and they cannot complete a computer task. Particularly, it is hypothesized that MS Office users act in a self-serving manner with respect to interface agents incorporated into the system:

H1. MS Office interface agent users attribute positive outcomes to their personal factors and negative outcomes to interface agents in the human–agent interaction process.

The psychology literature suggests that when a third party has a high degree of control over the critical outcome of a particular situation, people assign more external attributions because they feel the lack of ability to influence the course of events and consider themselves entirely dependent on others (Lee and Robinson, 2000; Lee and Tiedens, 2001). As interface agents become more autonomous, they possess more power and control over the outcomes of human–computer interaction processes. When individuals delegate autonomous tasks to interface agents, for example, the presentation of task-specific tips, they expect interface agents to understand user actions, to predict their needs, and to provide relevant, timely, and necessary suggestions. At the same time, when an agent helped a person complete a task by acting highly autonomously, it can be hypothesized that he or she will acknowledge the contribution of this agent. Recall MS Office interface agents offer

two kinds of user support: (1) help menu assistance that refers to a low degree of interface agent autonomy because a user initiates the human–agent interaction process and (2) real-time advice presentation that corresponds to a high degree of interface agent autonomy because an agent itself makes a decision when and how to offer user support. By following the reasoning above, it is hypothesized that:

H2. As the degree of autonomy of MS Office interface agents increases, users tend to assign more negative attributions to agents under the condition of failure and more positive attributions under the condition of success.

3. Methodology

In order to answer the study's research question and related hypotheses, an experiment with the actual and potential users of MS Office interface agents was conducted. The subjects were both undergraduate (third- and fourth-year B.Com.) and graduate (second-year MBA and Ph.D.) students of a North American university. The course settings of the programs require extensive usage of major MS Office applications, such as Word and Excel. The results of another independent survey of 243 students of the same university administered several months earlier by Serenko (2007) indicated that over 99% of these students were familiar with interface agents in MS Office applications. Moreover, 43% of them experimented with this technology by trying to personalize the settings of an agent. All computer laboratories of the school had interface agents installed for at least 3 years. Therefore, all students were knowledgeable enough about the usage of interface agents in MS Office.

In order to understand and to measure the self-serving biases of the subjects, individuals were asked to read four brief vignettes on the employment of interface agents in the MS Office suite. The application of vignettes in attribution theory experiments has a long-standing tradition (Weiner, 1980; Wadley and Haley, 2001). For instance, by utilizing vignettes in their investigations, Zeelenberg et al. (2000) analyzed people's attributions of responsibility and affective reactions to decision outcomes. Lee and Tiedens (2001) examined the influence of a person's status on possible attribution inferences. Vignettes are also employed in information technology investigations (Constant et al., 1994; Harrington, 1996; Gattiker and Kelley, 1999). For example, by using vignettes, Harrington (1995) analyzed ethical issues with respect to information systems personnel, and Jarvenpaa and Staples (2000) studied the use of collaborative electronic media for information sharing.

As discussed in the previous section, MS interface agents offer two distinct categories of user assistance: help menu support and real-time tips presentation. Therefore, two vignettes corresponding to each type of assistance were created. For each kind of assistance, one positive situation and one negative situation were presented. Therefore, subjects were offered four vignettes: Situation 1 – Positive

Outcome, Low Autonomy (Pos-LowAutonomy); Situation 2 – Positive Outcome, High Autonomy (Pos-High Autonomy); Situation 3 – Negative Outcome, Low Autonomy (Neg-LowAutonomy); and Situation 4 – Negative Outcome, High Autonomy (Neg-HighAutonomy).

Prior to administering the questionnaire to respondents, a pilot test of the instrument was conducted. The initial versions of vignettes were written by the author. A team comprised psychology, information systems, and human–computer interaction experts reviewed and commented on these vignettes in order to assess their face validity. Based on this feedback, text of the vignettes was modified until agreement was reached.

In order to create a research instrument accompanying these vignettes, the Attributional Style Questionnaire (ASQ) (Peterson et al., 1982) was adapted. Previous investigations demonstrate the viability of adapting the ASQ, depending on the field of application (Proudfoot et al., 2001). The purpose of the ASQ is to measure individual differences in the use of causes of a situation. The application of this instrument yields scores for individual differences in the tendencies to attribute the causes of good or bad events to internal (person-specific) or external (other people or circumstances-specific) factors. The ASQ consists of instructions, brief statements, and a measurement scale. The pilot test utilized four vignettes as well as the original instructions and adaptations of the ASQ questions. Appendix A offers the instrument.¹ It should be noted that the instrument presents both open-ended and close-ended items that allow triangulating qualitative and quantitative data. Triangulation is the combination of different methodologies to study the same phenomenon (Jick, 1979). The comparison of data obtained by different methods allows testing for external validity of the results and uncovering new findings that cannot be identified by a single method (Rohner, 1977; Creswell, 2003).

In the pilot study, the situations were presented in four different orders by following 4 × 4 Latin Square Design to control for order effects (Zeelenberg et al., 2000).² The following order was employed: S1–S2–S3–S4 (i.e., this is the order presented in the sample questionnaire, Appendix A), S2–S3–S4–S1, S3–S4–S1–S2, and S4–S1–S2–S3, and the pilot questionnaire was administered to a group of 30 people. A semi-structured instrument administration method was selected; when respondents had questions about instructions, situations, or scales, the researcher verbally provided all necessary clarifications. A brief interview with 15 respondents was conducted to solicit feedback on the instrument.

The results of the pilot study indicated that almost 50% of the respondents could not understand instructions as well as the question “*Please write down one major cause.*”

They suggested that questions should be more specific and that the term “cause” should be clearly explained. Subjects also indicated that the instructions were very long, redundant, and vague.³

A review of the methodologies of the original work that developed the ASQ (Peterson et al., 1982) as well as of a number of subsequent projects that utilized this instrument revealed that the ASQ was developed and validated through the involvement of psychology students in experiments. In contrast, the present investigation involved respondents who were not enrolled in psychology classes. Based on the comments of the respondents in the pilot study, it was suggested that the original ASQ cannot be applied to individuals who are not familiar with psychology terminology. Therefore, the instructions and questions were further adjusted to the level of comprehension of people who are not familiar with psychology research. An improved version of the questionnaire was reviewed by a team of experts. The results of the pilot study were excluded from further data analysis.

Similar to the pilot study, the situations in the full study were presented in four different orders, and the new instrument was administered to a group of 50 respondents who did not demonstrate any difficulty understanding and interpreting the instructions, vignettes, and questions. After that, 180 people were surveyed. In total, 230 surveys were administered and 228 questionnaires were collected. Sixteen of them were partially completed (despite the instructions, one or more items were left unanswered) and were excluded. Overall, 202 questionnaires were used for analysis.

4. Results

Sixty percent of the respondents were males and 40% were females. Their average age was 26, ranging from 20 to 50 years old. 71%, 48%, 72%, and 43% of the respondents stated that they experienced Situations 1, 2, 3, and 4, respectively, when they interacted with interface agents in the MS Office suite. Overall, 89% of the subjects indicated they came across at least one out of four situations presented in the vignettes.

In the first stage of the data analysis, the internal consistency of the scale measuring an attribution style of a person (questions 2 and 3) for each situation was calculated. The Cronbach’s α exceeded 0.8 for each situation. The attribution score for each situation was determined by the calculation of the average of questions 2 and 3 as suggested by Zeelenberg et al. (1998). Table 1 presents the means and standard deviations of attribution scores.⁴

³ In the pilot study, the original instructions of the Attributional Style Questionnaire were utilized. The instructions are presented in detail by Peterson et al. (1982).

⁴ Given that this study employs a repeated-measures design, no conclusions can be made based on differences in item means (i.e., paired comparisons are required).

¹ This appendix presents the final version of the instrument (i.e., the instrument was modified based on the results of the pilot study).

² In the pilot study, the effect of the order of situations was not investigated since the data were discarded for validity reasons.

Table 1
Attribution scores

Mean (std)	High autonomy	Low autonomy
Positive situation	3.57 (1.50)	4.21 (1.45)
Negative situation	3.60 (1.67)	4.04 (1.47)

In order to test the hypotheses, differences in attribution scores were compared. According to the methodology employed by attribution theory researchers, individuals are expected to assign different scores to causes of events depending on their attribution predispositions towards a particular situation. On a 7-point Likert-type scale, the difference in scores is expected to be statistically significant (Peterson et al., 1982). Therefore, attribution scores were analyzed in a 2 (Situation: Positive vs. Negative) \times 2 (Autonomy: High vs. Low) repeated-measures within-subjects ANOVA, and the order of situations was added as a variable. Based on the results, no interaction effects were observed between the order factor and the degree of autonomy, $F(3,198) = 0.512$, ns, and between the order factor and the situation, $F(3,198) = 1.244$, ns. This shows that the order in which the situations were presented did not have any statistically significant effect. In terms of H1, no main effect of the situation was identified, $F(1,198) = 0.393$, ns. With respect to H2, the main effect of the degree of autonomy was discovered, $F(1,198) = 36.543$, $p < 0.000$. No interaction effect between the type of a situation and the degree of autonomy was found, $F(1,198) = 1.341$, ns. Table 2 summarizes the results.

These findings were further confirmed by using a series of paired-samples t tests. If H1 holds true, each person's attribution score for a positive outcome situation should be higher than that for a similar negative outcome one as suggested by Peterson et al. (1982).

The following differences in the attribution scores were obtained in the present study:

$$\text{Pos-LowAutonomy} - \text{Neg-LowAutonomy} \\ = 0.17 (t = 1.197, p > 0.05, \text{ns})$$

$$\text{Pos-HighAutonomy} - \text{Neg-HighAutonomy} \\ = -0.03 (t = 0.31, p > 0.05, \text{ns})$$

Thus, H1 was rejected.

Table 2
Main and interaction effects

	F value	P value
Main effects		
Situation	0.393	ns
Autonomy	36.543	0.000
Interaction effects		
Order factor \times Autonomy	0.512	ns
Order factor \times Situation	1.244	ns
Situation \times Autonomy	1.341	ns

If H2 holds true, attribution scores in a negative-outcome situation which is characterized by a low degree of autonomy should be higher than those in a situation which is characterized by a higher degree of autonomy. Similarly, the same logic is applied to a positive-outcome situation:

$$\text{Neg-LowAutonomy} - \text{Neg-HighAutonomy} \\ = 0.44 (t = 3.292, p < 0.001)$$

$$\text{Pos-LowAutonomy} - \text{Pos-HighAutonomy} \\ = 0.64 (t = 6.208, p < 0.000)$$

Therefore, H2 was supported.

Recall in addition to the Likert-type scale items, the first open-ended question of the instrument asked respondents to provide one major cause of the situation. Two coders analyzed all open-ended responses by using the same codebook. The Krippendorff's agreement coefficient was 0.78 which falls into the acceptable range (Krippendorff, 1980). All discrepancies were discussed and mutual agreement was reached. The coders failed to classify or did not agree on the classification of only 4.3% of all items that further confirms the validity of the dataset. Figs. 1–4 present the reasons why respondents believed they achieved or failed to achieve their goals in the MS Office suite when they interacted with an interface agent.

The figures related to the positive-outcome situations indicate that people tend to give credit to an interface agent when it helped them complete a task successfully rather than to assign internal attributions to themselves. This empirical evidence contradicts the self-serving hypothesis of attribution theory. For example, in the Pos-LowAutonomy situation, 64% of the respondents attributed the cause of successful task completion to high-quality agent assistance (e.g., "I completed the task because the agent presented the information I needed"). At the same time, only 19% of the subjects referred to their strong personal abilities to use MS Office help (e.g., "I was good at using MS help," or "I typed in correct keywords"), 10% to their abilities to accept and utilize help from the agent ("I was able to use the agent's information"), 4% to their internal factors (e.g., "I concentrated" or "I was patient"), and 2% to their strong personal knowledge of MS Office applications (e.g., "Because I am good at using MS Word"). Several respondents indicated they did not believe that an interface agent may be useful, and they attributed the cause of the successful task completion to pure luck. The number of subjects who gave credit to an interface agent in the Pos-HighAutonomy situation (78%) was higher than that in the Pos-LowAutonomy situation (64%).

The figures related to the negative-outcome situations demonstrate that people do not always assign external attributions to interface agents. In the Neg-LowAutonomy situation, only 55% of respondents blamed an interface agent (e.g., "The agent did not present the information I needed"), whereas 42% of the subjects stated that it was their personal fault. As such, 20% attributed the cause of failure to their weak personal abilities to use an MS Office

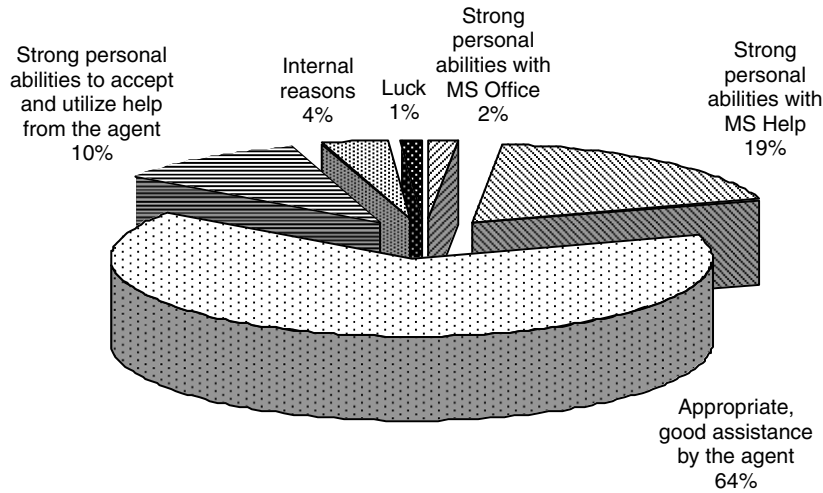


Fig. 1. Cause of Situation 1 (positive outcome, low degree of autonomy).

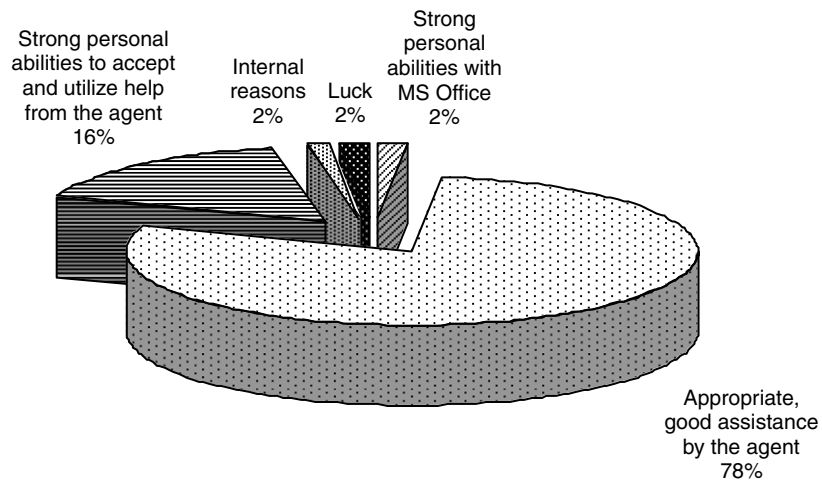


Fig. 2. Cause of Situation 2 (positive outcome, high degree of autonomy).

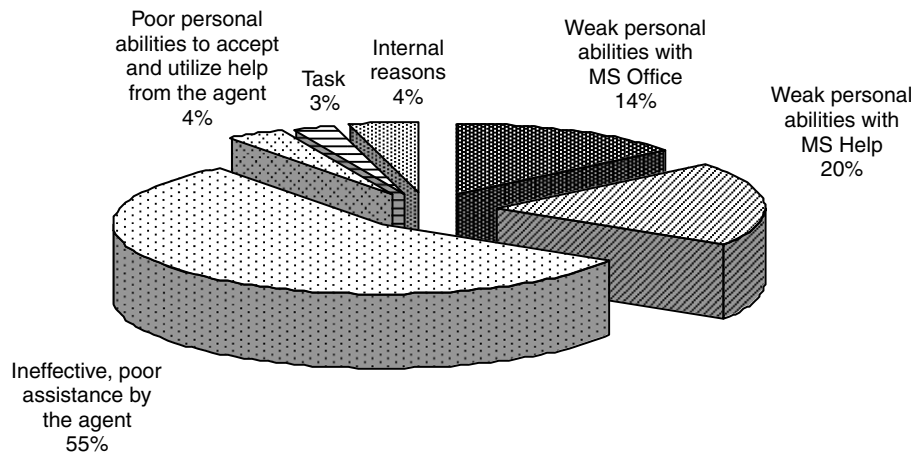


Fig. 3. Cause of Situation 3 (negative outcome, low degree of autonomy).

help system (e.g., “I was not good at using MS help” or “I entered wrong keywords”), 14% to their inadequate personal knowledge of MS Office (e.g., “Because I am not

good at using MS Office”), 4% to their internal reasons (“I did not concentrate enough” or “I was not patient”), and 4% to their inability to accept and utilize help from

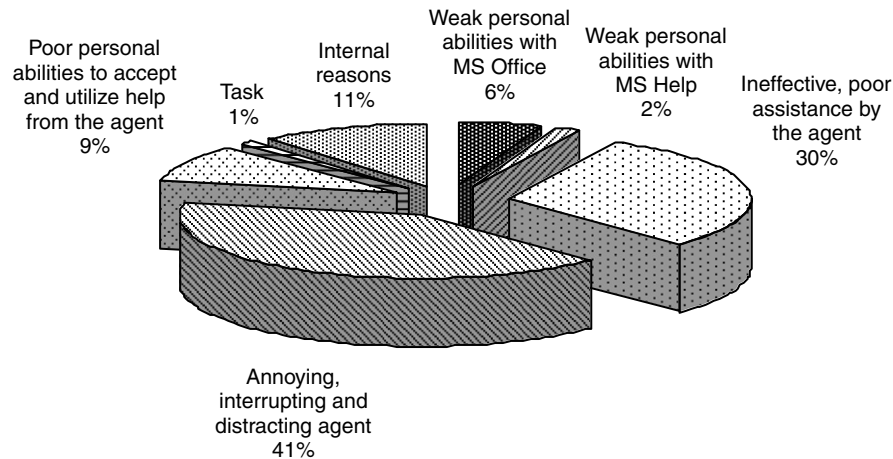


Fig. 4. Cause of Situation 4 (negative outcome, high degree of autonomy).

the agent (e.g., “I was not able to use the obtained information”). Several people mentioned the nature of the task (e.g., “The task was too difficult”). In the Neg-HighAutonomy situation, 41% of users mentioned an annoying, interrupting, or distracting agent (“The agent distracted me”). In total, in the Neg-HighAutonomy situation, negative external attributions comprised 71% of responses compared with only 55% in the Neg-LowAutonomy situation.

In order to further test the hypotheses, differences in the causes of attributions obtained through the open-ended question were compared between the situations by using the χ^2 test. This test was similar to the analysis conducted by comparing attribution scores obtained through a Likert-type scale. For each situation presented in Figs. 1–4, the number of causes attributed to an agent (i.e., external attributions) and to a user (i.e., internal attributions) was counted. After that, the differences in the number of categories between the situations were tested.

In terms of H1, the following differences were obtained:

$$\text{Pos-LowAutonomy} - \text{Neg-LowAutonomy} \rightarrow \chi^2(1) \\ = 2.18 (p > 0.05, \text{ns})$$

$$\text{Pos-HighAutonomy} - \text{Neg-HighAutonomy} \rightarrow \chi^2(1) \\ = 3.5 (p > 0.05, \text{ns})$$

With respect to H2, it was found that:

$$\text{Neg-LowAutonomy} - \text{Neg-HighAutonomy} \rightarrow \chi^2(1) \\ = 8.399 (p < 0.01)$$

$$\text{Pos-LowAutonomy} - \text{Pos-HighAutonomy} \rightarrow \chi^2(1) \\ = 10.609 (p < 0.01)$$

Again, H1 was rejected and H2 was supported. Therefore, it is concluded that the triangulation of the quantitative and qualitative questionnaire items demonstrates the validity of the findings. Overall, the results demonstrate that MS Office interface agent users do not always attribute the positive outcomes to their personal factors and negative

outcomes to interface agents in human–agent interaction. At the same time, as the level of an agent’s autonomy increases, users are willing to assign more positive attributions to interface agents under the condition of success and to hold interface agents more responsible for failures.

5. Discussion

The purpose of the study was to investigate the self-serving biases of interface agent users in the MS Office suite. The research question asked whether MS Office users attribute positive outcomes to their personal factors and negative outcomes to interface agents. There are three key points that deserve attention.

First, it is unarguable that MIS academics may dramatically benefit by adapting existing research techniques and instruments from reference disciplines (Straub, 1989). However, the author cautions that not every instrument may be successfully applied to a new field. Recall that the pilot study of this project utilized a part of the original Attributional Style Questionnaire that was previously found to be reliable and valid. Despite that, the study’s respondents expressed difficulty understanding and interpreting both instructions and questions because their educational background was different from that of subjects who were surveyed by the ASQ developers.⁵ Therefore, it is suggested that MIS researchers who intend to adapt existing instruments from reference disciplines always conduct a thorough reevaluation of an instrument’s reliability and validity.

Second, with respect to interface agents in the MS Office suite, self-serving bias was not observed (i.e., H1 was rejected). Instead, the empirical investigation demonstrated that people may attribute the cause of task success to an agent, and they may hold themselves responsible for task failure. This contradicts prior works by Moon (2003) and Moon and Nass (1998) who demonstrated the existence

⁵ Please refer to Section 3 for details on instrument adaptation.

of self-serving biases in human–computer interaction. At the same time, Ruvini and Gabriel (2002), who conducted user assessment of an interface agent for email, report that individuals were ready to tolerate a certain degree of errors produced by the intelligent assistant; this supports this study's findings.

Even though the existence of self-serving biases was not entirely supported, the present project does not reject the Computers as Social Actors Paradigm (Reeves and Nass, 1996; Nass and Moon, 2000; Nass, 2004); it is still believed that individuals respond to computer-based systems socially. Three assumptions may potentially explain the unexpected findings of this study.

The first proposition refers to the imperfection of the modern computer systems, including interface agents. Currently, many software applications present unnecessary complexity, unintuitive features, and limited functionality. As a result, individuals tend not to complain about software problems; they expect computers to be imperfect and attempt to learn how to deal with those shortcomings to improve their abilities to utilize the system. With respect to the MS Office suite, some users take an agent's failure to deliver what it is supposed to for granted whereas they are willing to praise it for success. A similar phenomenon was discovered by the developers of Aria, an agent for annotating and retrieving images (Lieberman et al., 2001). Agent testing demonstrated that some users were reluctant to show frustration about problems associated with the use of the agent because they expected computer software to be unreliable. Serenko (2006b) in his empirical investigation of user behavior towards interface agents also reports that some agent users did not show negative feelings towards an incident when an agent behaved highly unreliably.

The second explanation relates to the mitigating effect of an interpersonal context between a user and an interface agent. Previous psychology research suggests that when two people are engaged in intimate self-disclosure, they stop behaving in a self-serving manner (Campbell et al., 2000). Instead, they may have shared responsibility for successful and unsuccessful results. This line of reasoning may be also applied to computers. Recently, Moon (2003) demonstrated that when individuals have a long history of self-disclosure with a computer that leads to a high level of attraction towards the machine, they are less likely to blame the computer for negative outcomes and are more likely to credit the computer for positive ones. In terms of interface agents, it may be hypothesized that some users developed an intimate relationship with MS Office agents that reduced their self-serving bias. The third assumption refers to the potential differences in human-to-human and human-to-software agent interaction modes. For instance, it may be hypothesized that whereas people exhibit self-serving biases when they interact with their peers, they may not do so when they deal with artificial "intelligent" entities because the interaction environment is less familiar

and more ambiguous that leaves some room for shifting attributions in unconventional manners.

Third, when users perceive an interface agent as possessing a high degree of autonomy, they feel that the agent is responsible for delivering high-quality assistance. When the agent initiated communication with a user by offering a tip which was difficult to follow, most people attributed the negative outcome to the agent by calling it an "irritating, distracting, or useless agent." Agent designers should be aware that the more autonomous interface agents become, the more responsibility users will assign to agents if they fail to deliver what is expected. At the same time, when an agent that possesses a high degree of autonomy helps a person complete a computer-related task successfully, an individual is willing to acknowledge the contribution of the agent.

6. Limitations and conclusions

This study has several limitations. First, users of only one type of an interface agent were surveyed that constrains the generalizability of the findings. Currently, there are various categories of interface agents available; that would be interesting to see whether the conclusions of this study hold true with respect to other interface agents, for example, shopping or Web browsing agents. Second, this research shows how individuals make their attributions, but it does not explain why they behave this way; one may only hypothesize why users did not behave in a self-serving manner. It will also be interesting to test attribution theory with respect to agents that are invisible to the end users. For instance, if individuals know that there is an agent working in the background doing tasks for them but they do not directly interact with this entity, how will they attribute the causes of task successes and failures? This project does not offer an explanation. Third, several important moderating variables, such as computer expertise, innovativeness, task nature, and experience with MS Agents, in general, were omitted in this project. It is possible that they may potentially have an impact on attribution behaviors of a specific user group. Fourth, the vignettes utilized in this study's questionnaires pertained to two different applications (MS Word and MS Excel) but no application-specific effects were investigated. Future scholars may develop an instrument with multiple situations relating to only one application. In this case, they should control for common method bias to ensure that respondents actually distinguish among different situations and questionnaire items (Podsakoff and Organ, 1986; Podsakoff et al., 2003). Fifth, no distinction among MS Agent characters was made. Recall that MS Office users may select a specific interface agent in MS Office. One may enquire whether the look and feel of the agent affects a user's self-serving bias. It is hoped that future investigations will continue this line of inquiry and answer these questions. Overall, despite

these limitations, it is believed that this study was successful, and it may potentially improve our understanding of human–agent interaction processes.

The field of agent-based computing is in the early stages of development. The MS Office suite is one of the first end-user commercial products to include an interface agent. Most Office users have already experienced either a positive or a negative outcome of agent usage. It is suggested that developers consider this study’s findings and researchers continue investigating the self-serving biases of interface agent users.

Acknowledgments

The author is grateful to four anonymous reviewers who provided very valuable feedback on an earlier version of this paper. A very special thanks goes to Reviewer 2 for his/her suggestions on the statistical techniques employed in this project. This research was inspired by the study by Moon and Nass (1998).

Appendix A. Questionnaire

Definition: An MS Animated Agent is an interactive character that pops up when you use a help menu in a Microsoft application such as Word, Excel, PowerPoint, or Outlook. Sometimes, the agent pops up to present messages and tips intended for more efficient usage of these applications. Below are examples of MS Animated Agents:



Instructions: Please try to vividly imagine yourself in the situations that follow and answer the questions as if such situations happened to you.

General story: You are an employee of a large company. Your work includes the preparation of reports of the department. You spend most of your time working with MS Office applications.

Situation 1. Once you wanted to find out how to prevent a table from breaking across pages in MS Word. You selected “Help” on the menu, and an animated agent popped up. You typed a few keywords, and the agent provided the exact information you looked for. This helped you place the tables the way you wanted to. You were very happy with a nice looking report.

1. Write down the *one* major cause of the successful task completion (i.e., one major reason why this task was completed successfully)_____

2. Is the cause of the successful completion of the task due to something about you or to something about the animated agent? (circle one number)

totally due to the agent		totally due to me
1	2	3
		4
		5
		6
		7

3. Who should be credited for the successful task completion? (circle one number)

totally the agent		totally me
1	2	3
		4
		5
		6
		7

4. Have you ever experienced a similar situation?
Yes No

Situation 2. When you were concentrating on a difficult task with MS Excel, an animated agent popped up. It presented a tip on a more efficient use of keyboard shortcuts. You followed the tip and completed the task more efficiently than before.

1. Write down the *one* major cause of the efficient task completion (i.e., one major reason why this task was completed efficiently)_____

2. Is the cause of the efficient task completion due to something about you or to something about the animated agent? (circle one number)

totally due to the agent		totally due to me
1	2	3
		4
		5
		6
		7

3. Who should be credited for the efficient task completion? (circle one number)

totally the animated agent		totally me
1	2	3
		4
		5
		6
		7

4. Have you ever experienced a similar situation?
Yes No

Situation 3. Once you wanted to find out how to prevent a table from breaking across pages in MS Word. You selected “Help” on the menu, and an animated agent popped up. You tried to type different keywords, but the agent did not offer any useful information. You spent 10 min trying different keywords, but did not find what you looked for. You clicked “Hide,” and the agent disappeared. You finished the document with tables spanning across multiple pages. You were very disappointed by the look of the report.

1. Write down the *one* major cause of the unsuccessful task completion (i.e., one major reason why this task was completed unsuccessfully)_____

2. Is the cause of the unsuccessful task completion due to something about you or to something about the animated agent? (circle one number)

totally due to the agent		totally due to me
1	2	3
		4
		5
		6
		7

3. Who is responsible for the unsuccessful task completion? (circle one number)

totally the agent 1 2 3 4 5 6 7 totally me

4. Have you ever experienced a similar situation? Yes No

Situation 4. When you were concentrating on a difficult task with MS Excel, an animated agent popped up. It presented a tip on a more efficient use of keyboard shortcuts. You read the tip, but found it complicated and declined it. After that, you realized you lost track of your previous activity and you had to go back to re-do some work.

1. Write down the one major cause of the inefficient task completion (i.e., one major reason why this task was completed inefficiently)

2. Is the cause of the inefficient task completion due to something about you or to something about the animated agent? (circle one number)

totally due to the agent 1 2 3 4 5 6 7 totally due to me

3. Who is responsible for the inefficient task completion? (circle one number)

totally the agent 1 2 3 4 5 6 7 totally me

4. Have you ever experienced a similar situation? Yes No

Demographic information:

Your age: _____ (years)

Your gender: male female

Other comments (if any)

References

Ball, G., Ling, D., Kurlander, D., Miller, J., Pugh, D., Skelly, T., Stankosky, A., Thiel, D., Dantzich, M.V., Wax, T., 1997. Lifelike computer characters: the Persona project at Microsoft Research. In: Bradshaw, J.M. (Ed.), Software Agents. The AAAI Press/The MIT Press, Menlo Park, CA, pp. 191–222.
Blackwood, N.J., Bentall, R.P., Ffytche, D.H., Simmons, A., Murray, R.M., Howard, R.J., 2003. Self-responsibility and the self-serving bias: an fMRI investigation of causal attributions. NeuroImage 20 (2), 1076–1085.
Campbell, W.K., Sedikides, C., Reeder, G.D., Elliot, A.J., 2000. Among friends? An examination of friendship and the self-serving bias. Br. J. Soc. Psychol. 39 (2), 229–239.
Constant, D., Kiesler, S., Sproull, L., 1994. What’s mine is ours, or is it? A study of attitudes about information sharing. Inf. Syst. Res. 5 (4), 400–421.

Creswell, J.W., 2003. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage Publications, Thousand Oaks, CA.
Detlor, B., 2004. Towards Knowledge Portals: From Human Issues to Intelligent Agents. Kluwer Academic Publishers, Dordrecht, The Netherlands.
Farmer, A., Pecorino, P., 2002. Pretrial bargaining with self-serving bias and asymmetric information. J. Econ. Behav. Organ. 48 (2), 163–176.
Friedman, B., 1995. ‘It’s the computer’s fault’: reasoning about computers as moral agents. In: Proceedings of the Conference on Human Factors in Computing Systems, Denver, Colorado. The ACM Press, New York.
Friedman, B. (Ed.), 1997. Human Values and the Design of Computer Technology. Cambridge University Press, Cambridge, MA.
Gadhoom, Y., 1999. Potential effects of managers’ entrenchment and shareholdings on competitiveness. Eur. J. Oper. Res. 118 (2), 332–349.
Gattiker, U.E., Kelley, H., 1999. Morality and computers: attitudes and differences in judgments. Inf. Syst. Res. 10 (3), 233–254.
Harrington, S.J., 1995. The anomaly of other-directedness: when normally ethical I.S. personnel are unethical. Comput. Pers. 16 (2), 3–11.
Harrington, S.J., 1996. The effect of codes of ethics and personal denial of responsibility on computer abuse judgments and intentions. MIS Q. 20 (3), 257–278.
Heider, F., 1958. The Psychology of Interpersonal Relations. Wiley, New York.
Jarvenpaa, S.L., Staples, S., 2000. The use of collaborative electronic media for information sharing: an exploratory study of determinants. J. Strateg. Inf. Syst. 9 (2–3), 129–154.
Jick, T.D., 1979. Mixing qualitative and quantitative methods: triangulation in action. Adm. Sci. Q. 24 (4), 602–611.
Johnson, W.L., Rickel, J.W., Lester, J., 2000. Animated pedagogical agents: face-to-face interaction in interactive learning environments. Int. J. Artif. Intell. Educ. 11, 47–78.
Jones, E.E., Davis, K., 1965. From acts to dispositions: the attribution process in person perception. In: Berkowitz, L. (Ed.), Advances in Experimental Social Psychology. Academic Press, New York, vol. 2, pp. 219–266.
Keeble, R.J., Macredie, R.D., 2000. Assistant agents for the world wide web intelligent interface design challenges. Interact. Comput. 12 (4), 357–381.
Kelly, H.H., 1972. Attribution in social interaction. In: Jones, E.E., Kanouse, D.E., Nelly, H.H., et al. (Eds.), Attribution: Perceiving the Causes of Behavior. General Learning Press, Morristown, NJ, pp. 1–27.
Kling, R. (Ed.), 1996. Computerization and Controversy: Value Conflicts and Social Choices. Academic Press, San Diego.
Krippendorff, K., 1980. Content Analysis: An Introduction to its Methodology. Sage Publications, Beverly Hills, CA.
Lashkari, Y., Metral, M., Maes, P., 1994. Collaborative interface agents. In: Proceedings of the Twelfth National Conference on Artificial Intelligence, Seattle, Washington. AAAI Press, Menlo Park, CA.
Lee, F., Robinson, R., 2000. An attributional analysis of social accounts: implications of playing the blame game. J. Appl. Soc. Psychol. 30 (9), 1853–1879.
Lee, F., Tiedens, L.Z., 2001. Who’s being served? ‘Self-Serving’ attributions in social hierarchies. Organ. Behav. Hum. Decis. Process. 84 (2), 254–287.
Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhoga, R.S., 1997. The persona effect: affective impact of animated pedagogical agents. In: Proceedings of the Conference on Human Factors in Computing Systems, Atlanta, Georgia. The ACM Press, New York.
Lieberman, H., 1995. Letizia: an agent that assists Web browsing. In: Proceedings of the International Joint Conference on Artificial Intelligence, Montreal, Canada.
Lieberman, H., 2001. Interfaces that give and take advice. In: Carroll, J.M. (Ed.), Human-Computer Interaction in the New Millennium. The ACM Press, New York, pp. 475–484.

- Lieberman, H., Rosenzweig, E., Singh, P., 2001. Aria: an agent for annotating and retrieving images. *IEEE Comput.* 34 (7), 57–62.
- Lieberman, H., Selker, T., 2003. Agents for the user interface. In: Bradshaw, J.M. (Ed.), *Handbook of Agent Technology*. The MIT Press, Cambridge, MA.
- Maes, P., 1994. Agents that reduce work and information overload. *Commun. ACM* 37 (7), 31–40.
- Maes, P., 1995. Artificial life meets entertainment: lifelike autonomous agents. *Commun. ACM* 38 (11), 108–114.
- Maes, P., Kozierok, R., 1993. Learning interface agents. In: *Proceedings of the Eleventh National Conference on Artificial Intelligence*, Washington, DC. The MIT Press, Cambridge, MA.
- Magid, L.J., 2001. Office XP – Easier to Use – No Wacky Paper Clip. *Los Angeles Times*. Available at http://www.pcanswer.com/articles/lat_officexp.htm. Retrieved December 10, 2003.
- Miller, D.T., Ross, M., 1975. Self-serving biases in the attribution of causality: facts or fiction? *Psychol. Bull.* 82 (2), 213–225.
- Moon, Y., 2003. Don't blame the computer: when self-disclosure moderates the self-serving bias. *J. Consum. Psychol.* 13 (1–2), 125–137.
- Moon, Y., Nass, C., 1998. Are computers scapegoats? Attributions of responsibility in human–computer interaction. *Int. J. Hum.–Comput. Stud.* 49 (1), 79–94.
- Nass, C., 2004. Etiquette equality: exhibitions and expectations of computer politeness. *Commun. ACM* 47 (4), 35–37.
- Nass, C., Fogg, B.J., Moon, Y., 1996. Can computers be teammates? *Int. J. Hum. Comput. Stud.* 45 (6), 669–678.
- Nass, C., Moon, Y., 2000. Machines and mindlessness: social responses to computers. *J. Soc. Iss.* 56 (1), 81–103.
- Nass, C., Steuer, J., Henriksen, L., Dryer, D.C., 1994. Machines, social attributions, and ethopoeia: performance assessments of computers subsequent to 'self-' or 'other-' evaluations. *Int. J. Hum. Comput. Stud.* 40 (3), 543–559.
- Person, N.K., Craig, C., Price, P., Hu, X., Gholson, B., Greasser, A.C., 2000. Incorporating human-like conversational behaviors into Auto-Tutor. In: *Proceedings of the Workshop on Achieving Human-like Behavior in the Interactive Animated Agents at the Agents 2000 Conference*, Barcelona, Catalonia, Spain. The ACM Press, New York.
- Peterson, C., Semmel, A., von Baeyer, C., Abramson, L.Y., Metalsky, G.I., Seligman, M.E.P., 1982. The Attributional Style Questionnaire. *Cogn. Ther. Res.* 6 (3), 287–300.
- Peterson, D.K., Kim, C., Kim, J.H., Tamura, T., 2002. The perceptions of information systems designers from the United States, Japan, and Korea on success and failure factors. *Int. J. Inf. Manage.* 22 (6), 421–439.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y., Podsakoff, N.P., 2003. Common method biases in behavioral research: a critical review of the literature and recommended remedies. *J. Appl. Psychol.* 88 (5), 879–903.
- Podsakoff, P.M., Organ, D.W., 1986. Self-report in organizational research: problems and prospects. *J. Manage.* 12 (4), 531–544.
- Postman, N., 1992. *Technopoly: The Surrender of Culture to Technology*. Alfred A. Knopf, New York.
- Proudfoot, J.G., Corr, P.J., Guest, D.E., Gray, J.A., 2001. The development and evaluation of a scale to measure occupational attributional style in the financial services sector. *Pers. Individ. Differences* 30 (2), 259–270.
- Reeves, B., Nass, C., 1996. *The Media Equation: How People Treat Computers, Televisions, and New Media Like Real People and Places*. CSLI Publications, Stanford, CA.
- Roch, S.G., Lane, J.A.S., Samuelson, C.D., Allison, S.T., Dent, J.L., 2000. Cognitive load and the equality heuristic: a two-stage model of resource overconsumption in small groups. *Organ. Behav. Hum. Decis. Process.* 83 (2), 185–212.
- Rohner, R.P., 1977. Advantages of the comparative method of anthropology. *Behavior Science Research* 12, 117–144.
- Ruvini, J.-D., Gabriel, J.-M., 2002. Do users tolerate errors from their assistant? Experiments with an E-mail classifier. In: *Proceedings of the Seventh International Conference on Intelligent User Interfaces*, San Francisco, CA.
- Schwenk, C., 1993. Management tenure and explanations for success and failure. *Omega* 21 (4), 449–456.
- Serenko, A., 2006a. The importance of interface agent characteristics from the end-user perspective. *Int. J. Intell. Inf. Technol.* 2 (2), 48–59.
- Serenko, A., 2006b. The use of interface agents for email notification in critical incidents. *Int. J. Hum. Comput. Stud.*, doi:10.1016/j.ijhcs.2006.06.004.
- Serenko, A., 2007. The development of an instrument to measure the degree of animation predisposition of agent users. *Comput. Hum. Behav.* 23 (1), 478–495.
- Serenko, A., Detlor, B., 2004. Intelligent agents as innovations. *AI Soc.* 18 (4), 364–381.
- Shneiderman, B., Maes, P., 1997. Direct manipulation vs. interface agents. *Interactions* 4 (6), 42–61.
- Straub, D.W., 1989. Validating instruments in MIS research. *MIS Q.* 13 (2), 147–169.
- Tennen, H., Affleck, G., 1990. Blaming others for threatening events. *Psychol. Bull.* 108 (2), 209–232.
- Trott, B., 1998. Microsoft's paper-clip assistant killed in Denver, CNN News. Available online at www.cnn.com/TECH/computing/9810/16/clipdeath.idg. Retrieved December 3, 2003.
- Wadley, V.G., Haley, W.E., 2001. Diagnostic attributions versus labeling: impact of Alzheimer's disease and major depression diagnoses on emotions, beliefs, and helping intentions of family members. *J. Gerontol. Psychol. Sci.* 56B (4), 244–252.
- Weiner, B., 1980. A cognitive (attribution)-emotion-action model of helping behavior: an analysis of judgments of help giving. *J. Pers. Soc. Psychol.* 39, 186–200.
- Weiner, B., Amirkhan, J., Folkes, V., Verette, J., 1987. An attributional analysis of excuse giving: studies of a naïve theory of emotion. *J. Pers. Soc. Psychol.* 52 (2), 316–324.
- Weiner, B., Frieze, I., Kukla, A., Reed, L., Rest, S., Rosenbaum, R.M., 1972. Perceiving the causes of success and failure. In: Jones, E.E., Kanouse, D.E., Nelly, H.H., et al. (Eds.), *Attribution: Perceiving the Causes of Behavior*. General Learning Press, Morristown, NJ, pp. 95–120.
- Zeelenberg, M., van der Pligt, J., de Vries, N.K., 2000. Attributions of responsibility and affective reactions to decision outcomes. *Acta Psychol.* 104 (3), 303–315.
- Zeelenberg, M., van Dijk, W.W., van der Pligt, J., Manstead, A.S.R., van Empelen, P., Reinderman, D., 1998. Emotional reactions to the outcomes of decisions: the role of counterfactual thought in the experience of regret and disappointment. *Organ. Behav. Hum. Decis. Process.* 75 (2), 117–141.