Persuasive Effects of Embodied Conversational Agent Teams

Hien Nguyen, Judith Masthoff, Pete Edwards

Computing Science Department, University of Aberdeen, UK
{hnguyen, jmasthoff, pedwards}@csd.abdn.ac.uk

Abstract. In a persuasive communication, not only the content of the message but also its source, and the type of communication can influence its persuasiveness on the audience. This paper compares the effects on the audience of direct versus indirect communication, one-sided versus two-sided messages, and one agent presenting the message versus a team presenting the message.

1 Introduction

Persuasive communication is "any message that is intended to shape, reinforce or change the responses of another or others." [11]. In other words, in a persuasive communication, a source attempts to influence a receiver's attitudes or behaviours through the use of messages. Each of these three components (the source, the receiver, and the messages) affects the effectiveness of persuasive communication. In addition, social psychology suggests that the *type of communication* (e.g. direct versus indirect) can impact a message's effectiveness [17,5].

The three most recognised characteristics of the source that influence its persuasiveness are *perceived credibility*, *likeability* and *similarity* [14,17]. These are not commodities that the source possesses, but they are the receiver's perception about the source. *Appearance cues* of the source (e.g. a white lab coat can make one a doctor) have been shown to affect its perceived credibility [17]. Hence, there has been a growing interest to use Embodied Conversational Agents (ECAs) in persuasive systems, and to make ECAs more persuasive.

In this paper, we explore persuasive ECAs in a healthcare counseling domain. More and more people use the Internet to seek out health related information [15]. Hence, automated systems on the Internet have the potential to provide users with an equivalence of the "ideal" one-on-one, personalised interaction with an expert to adopt health promoting behaviour more economically and conveniently. Bickmore argued that even if automated systems are less effective than actual one-on-one counselling, they still result in greater impact due to their ability to reach more users (impact = efficacy x reach) [4]. A considerable amount of research has been devoted to improve the efficacy of such systems, most of which focused on personalised content generation with various levels of personalisation [18]. Since one important goal of such systems is to persuade their users to adopt new behaviours, it is also vital

that they can win trust and credibility from users [17].In this paper, the ECAs will be fitness instructors, trying to convince users to exercise regularly.

Our research explores new methods to make automated health behaviour change systems more persuasive using personalised arguments and a team of animated agents for information presentation. In this paper, we seek answers to the following questions:

- *RQ1:* Which type of communication supports the persuasive message to have more impact on the user: *indirect* or *direct*? In the indirect setting, the user obtains the information by following a conversation between two agents: an instructor and a fictional character that is similar to the user. In the direct setting, the instructor agent gives the information to the user directly.
- *RQ2:* Does the use of a team of agents to present the message make it more persuasive than that of a single agent? In the former setting, each agent delivers a part of the message. In the latter setting, one agent delivers the whole message.
- *RQ3:* Is a two-sided message (a message that discusses the pros and cons of a topic) more persuasive than a one-sided message (a message that discusses only the pros of the topic)?

2 Related Work

Animated characters have been acknowledged to have positive effects on the users' attitudes and experience of interaction [7].

With respect to the persuasive effect of adding social presence of the source, mixed results have been found. Adding a formal photograph of an author has been shown to improve the trustworthiness, believability, perceived expertise and competence of a web article (compared to an informal or no photograph) [9]. However, adding an image of a person did not increase the perceived trustworthiness of a computer's recommendation system [19]. It has been suggested that a photo can boost trust in e-commerce websites, but can also damage it [16]. A series of studies found a positive influence of the similarity of human-like agents to subjects (in terms of e.g. gender and ethnicity) on credibility of a teacher agent and motivation for learning (e.g. [3]). Our own work also indicated that the source's appearance can influence his/her perceived credibility, and prominently showing an image of a highly credible source with respect to the topic discussed in the message can have a positive effect on the message's perceived credibility, but that of a lowly credible source can have an opposite effect [13].

With respect to the use of a team of agents to present information, Andre et al [1] suggested that a team of animated agents could be used to reinforce the users' beliefs by allowing us to repeat the same information by employing each agent to convey it in a different way. This is in line with studies in psychology, which showed the positive effects of social norms on persuasion (e.g. [8,12]).

With respect to the effect of different communication settings, Craig et al showed the effectiveness of indirect interaction (where the user listens to a conversation between two virtual characters) over direct interaction (where the user converses with the virtual character) in the domain of e-learning [6]. In their experiment, users

significantly asked more questions and memorized more information after listening to a dialogue between two virtual characters. We can argue that in many situations, particularly when we are unsure about our position on a certain topic, we prefer hearing a conversation between people who have opposite points of view on the topic to actually discussing it with someone. Social psychology suggests that in such situations, we find the sources more credible since we think they are not trying to persuade us (e.g. [2]).

2 Experiment 1

2.1 Experimental Design

The aim of this experiment is to explore the questions raised in Section 1. To avoid any negative effect of the lack of realism of virtual characters' animation and voice using Text-To-Speech, we implement our characters as static images of real people with no animation or sound. The images of the fitness instructors used have been verified to have high credibility with respect to giving advice on fitness programmes in a previous experiment [10].

Forty-one participants took part in the experiment (mean age = 26.3; stDev = 8.4, predominately male). All were students on an HCI course in a university Computer Science department. Participants were told about a fictional user John, who finds regular exercise too difficult, because it would prevent him from spending time with friends and family (extremely important to him), there is too much he would need to learn to do it (quite important) and he would feel embarrassed if people saw him do it (quite important).

Participants were shown a sequence of screens, showing the interaction between John and a persuasive system about exercising. The experiment used a between subject design: participants experienced one of four experimental conditions, each showing a different system (see Table 1 for example screenshots):

- *C1: two-sided, indirect, one agent.* The interaction is indirect: John sees a conversation between fitness instructor Christine and Adam, who expresses similar difficulties with exercising to John. Christine delivers a two-sided message: for each reason that Adam mentions, Christine acknowledges it, gives a solution, and then mentions a positive effect of exercise.
- *C2: two-sided, direct, one agent.* The interaction is direct: Christine addresses John directly. Christine delivers the same two-sided message as in Condition C1.
- *C3: one-sided, direct, one agent.* The interaction is direct. However, Christine only delivers a one-sided message. She acknowledges the difficulties John has, but does not give any solution. She mentions the same positive effects of exercise as in Conditions C1 and C2.
- *C4: one-sided, direct, multiple agents.* The interaction is direct and the message one-sided. However, the message is delivered by three instructors instead of one: each instructor delivers a part of it, after saying they agreed with the previous instructor. The message overall is the same as in Condition C3.

Table 1. Examples of the screens shown to the participants in each condition

C1: two-sided, indirect, one agent	Most importantly, I feel that exercising will prevent me from spending time with my friends and family.   Well, you can make it part of your daily routine. For example, try to take the stairs instead of the elevator, or attend 30-minute lunch time routines in the gym. You would have a lot more energy for your friends and family if you exercised regularly.
C2: two-sided, direct, one agent	 Firstly, it doesn't take as much time as you think. It won't prevent you from spending time with your family and friends. You can make it part of your daily routine. For example, try to take the stairs instead of the elevator, or attend 30-minute lunch time routines in the gym. You would have a lot more energy for your friends and family if you exercised regularly.
C3: one-sided, direct, one agent	 I know that you want to exercise more often, but you also have many reasons for not doing it. But let's look at the positive side. You would have a lot more energy for your friends and family if you exercised regularly.
C4: one-sided, direct, multiple agents	 Christine is right. Besides, you would feel less stressed and have a better mood for the rest of the day. 

A comparison between conditions C1 and C2 will explore research question RQ1: whether direct or indirect messages work better. A comparison between C2 and C3 will explore RQ3: whether one- or two-sided messages work better. Finally, a comparison between C3 and C4 will explore RQ2: whether messages work better with one agent as source or multiple agents.

We decided to ask participants not only about the system's likely impact on opinion change, but also how easy to follow the system is, and how much they enjoyed it. In an experimental situation, participants are more likely to pay close attention to a system, and put effort into understanding what is going on. In a real situation, a user may well abandon a system if they find it too difficult to follow, and pay less attention to the message if they get bored. Previous research has indeed shown that usability has a high impact on the credibility of a website [10]. So, enjoyment and understandability are contributing factors to persuasiveness, which participants may well ignore due to the experimental situation, and are therefore good to measure separately. Participants answered three questions on a 7-point Likert scale:

- How easy to follow did you find the site? (from "very difficult" to "very easy"),
- How boring did you find the site? (from "not boring" to "very boring"), and
- Do you think a user resembling John would change his/her opinion on exercise after visiting this site? (from "not at all" to "a lot").

They were also asked to explain their answer to the last question.

2.2 Results and Discussion

Figure 1 shows results for each condition and each question. With respect to the likely impact on changing a user's opinion about exercise, a one-way ANOVA test indicated that there is indeed a difference between the four conditions ($p < 0.05$). Comparing each pair of conditions, we found a significant difference between each of C1, C2, C3 on the one hand and C4 on the other ($p<0.05$ for all). Participants' comments confirmed that they thought the multiple agents condition less persuasive, with some mentioning that the use of multiple instructors makes the system almost hectoring in tone and a bit patronizing. The difference between C1 and C2 was not significant, but the trend is for indirect messages to be more persuasive. The difference between C2 and C3 is not significant, but the trend is for two-sided messages to be more persuasive.

With respect to how easy each system is to follow, all four systems scored well on the scale, on average ranging from 5.1 to 6.5 out of 7. This indicates that the users have no difficulty in using any of these alternative, dialog-based user interfaces. Two-sided conditions C1 and C2 were rated on average as easier to follow than one-sided conditions C3 and C4, but only the difference between C2 and C3 was significant ($p<0.05$).

With respect to boredom of each system, two-sided conditions C1 and C2 were rated on average as less boring than one-sided conditions C3 and C4, but this difference was not significant. However, there was a significant correlation between boredom and opinion change (Pearson's correlation = -0.431, $p<0.01$): the more boring the participants found the system, the less impact they thought it would have on the user.

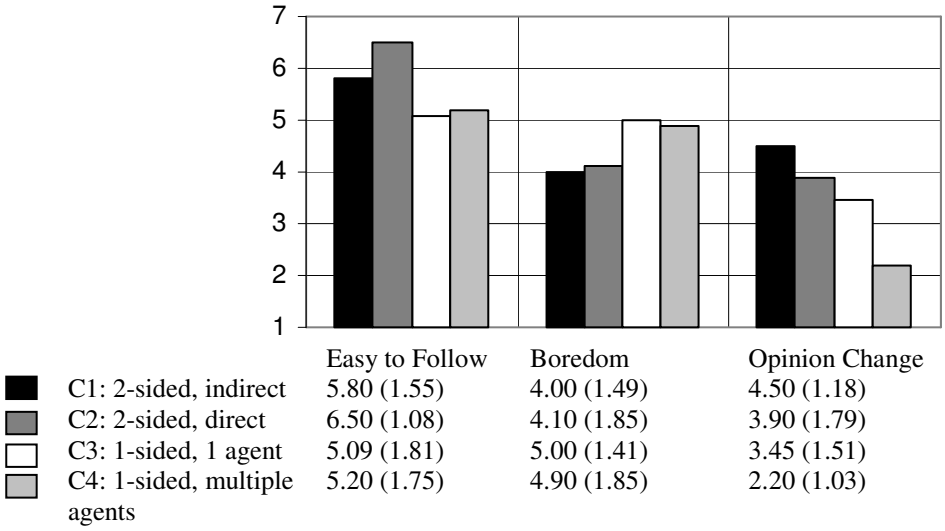

Fig. 1. The average values and standard deviations of each criterion for each group

In summary, we conclude the following regarding our research questions:

1. We cannot yet conclude whether indirect communication is more persuasive than direct communication when comparing conditions C1 and C2. The trend is in favour of indirect communication.
2. Surprisingly, the use of a team of agents to present information in this experiment considerably damaged the impact of the persuasive message. Perhaps the fact that all the agents appeared on screen at the same time, made the participants feel hectored and patronised as mentioned by some. An alternative would be to have each agent appear at a different stage of the conversation, re-emphasize what other agents have said and add new supporting information. We may also want to reconsider the way in which one agent supports what another has just said. Currently, this happens via sentences such as "Christine is right". Perhaps it would be better to instead reword the argument Christine has given.
3. The trend in the data suggests that two-sided messages are more persuasive than one-sided ones. This is supported by the two-sided message in C2 being significantly easier to follow than the corresponding one-sided message (C3). It is also supported by the trend in the boredom data, with C2 seeming to be perceived as less boring than C3.

3 Experiment 2

3.1 Experimental Design

The aim of this experiment is to extend and provide an alternative approach to our first experiment. In the previous experiment, participants looked at screenshots of a system. Clearly, this is not as engaging as actually experiencing the system. In this

experiment, we tried to give participants a more realistic experience. Instead of screenshots, they watched videos of the system in action. In the previous experiment, we used a between-subject design, and participants only saw one system. In this experiment, we used a within-subject design, with participants comparing between systems, hoping this would lead to a clearer view of what users prefer and regard as more persuasive. We used the same four system variants as in the first experiment, but also added another two, which used only text and were without any agents:

- *C5: two-sided, direct, text only.* This shows the text of the message used in C2.
- *C6: one-sided, direct, text only.* This shows the text of the message used in C3.

The new variants are intended to act as baselines.

Participants were asked to watch all six videos. The user interface was designed such that they could watch each video as many times as they liked. The order of the videos was randomized. Participants were asked to sort the six videos in order of (1) their likely impact on John's attitudes towards exercise, (2) how easy is system is to follow, and (3) boredom. The following section discusses our preliminary results on the basis of thirteen participants (mean age = 28, stDev = 7.7; 4 males, 9 females).

3.2 Results and Discussion

Participants were largely consistent in their judgements on the impact of each system on John's attitudes towards exercise and boredom (Cronbach's alpha=0.58, $p<0.05$, and alpha=.81, $p<0.01$, respectively). However, the same result did not hold for their judgement on how easy to follow each system is.

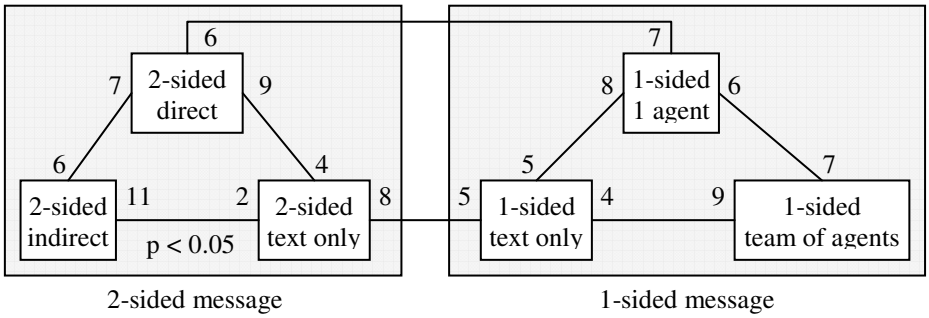

Fig. 2. Number of participants who preferred each system on persuasiveness (Experiment 2).

Figure 2 shows the participants' opinion on persuasiveness when comparing systems pair-wise. Numbers on the line between two systems show how many participants preferred each system. For instance, nine participants preferred the 2-sided, direct system to the 2-sided, text-only system, while four preferred the latter. The trend is for all systems with the agent(s) visible to be perceived as more convincing than their text-only counterparts. However, with the current limited number of participants, only the difference between the two-sided indirect condition and its text-only counterpart has reached statistical significance ($p<0.05$). In Experiment 1, we found trends in favour of indirect messages and 2-sided messages.

We do not find these here: participants were evenly divided. Even more surprisingly, in this experiment, the participants liked condition C4 as much as C1, C2, and C3, whilst in Experiment 1, we found a clear, statistically significant difference, with C4 being liked less. This raises the question of whether participants are actually capable of performing an accurate comparison in persuasiveness when given multiple systems to compare. Perhaps this task makes participants think too much about the differences between systems, while persuasion often takes a peripheral route as well. It may well be that a between-subject design, as in Experiment 1, is more appropriate.

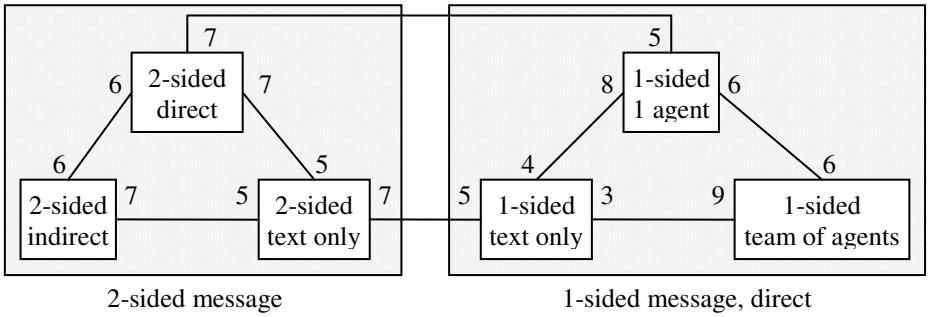

Fig. 3. Number of participants who preferred each system on how easy it is to follow (Exp. 2)

Figure 3 shows the participants' preference on how easy to follow systems are when comparing them pair-wise[1]. Whilst in Experiment 1 we found a trend for 2-sided messages being easier to follow than one-sided messages (with a statistically significant difference between C2 and C3), we do not find a similar result here. There is only a slight trend for 2-sided text only messages being easier to follow than 1-sided text only messages. However, there is a trend that the 1-sided text-only message is harder to follow than 1-sided messages with 1 or more agents.

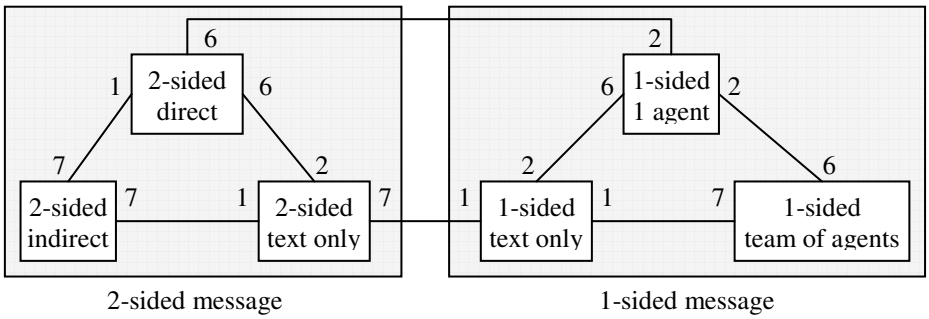

Fig. 4. Number of participants who preferred each system on avoidance of boredom (Exp. 2)

Figure 4 shows the participants' preference on avoidance of boredom when comparing systems pair-wise[2]. As in Experiment 1, we again found a trend for 2-sided

[1] One participant rated all videos as equally easy to follow, so is not included in Figure 3.

[2] Five participants did not complete the task correctly (e.g. one video was rated more than once), and their results are not included in Figure 4.

messages being less boring than one-sided messages. We also found a trend for the indirect 2-sided message being preferred to the direct one, and the team of agents being preferred to one agent. Furthermore, all the systems with agent(s) visible were perceived to be less boring than their text-only counterparts. Although all these differences were not statistically significant (non-surprising given the small number of participants), we believe that they will reach significance once we have more participants, given how clear the trends are.

In summary, from this experiment, we conclude the following:

1. There was no clear evidence that indirect communication is more persuasive than direct communication or vice versa when comparing conditions C1 and C2, although participants seemed to perceive indirect communication to be less boring.
2. There was also no significant preference between the use of one agent and a team of agents to present information.
3. With respect to the persuasiveness of a two-sided message compared to that of a one-sided message, similarly no clear preference was found, although participants seemed to perceive two-sided messages to be less boring.

The lack of results may well be because this experiment was too difficult. Some participants complained that they found it hard to remember and compare six videos. The number not finishing the final question also points in this direction. As mentioned above, a between-subject design may also be more appropriate.

4 Conclusions

This paper has investigated the persuasive effects on the audience's attitudes of direct versus indirect communication, one-sided versus two-sided messages, and one agent presenting the message versus a team presenting the message. Our second experiment suggests that dialog-based systems with the visual appearance of a conversational agent(s) are preferred over systems that use text only, as they are perceived to be more personal and caring, less boring, and to some extent easier to follow. When comparing our four dialog-based systems, we found somewhat conflicting results. Experiment 1 suggested a clear trend in which a two-sided message presented in an indirect communication was the most persuasive, followed by a two-sided message presented in a direct communication, a one-sided message presented by one agent, and a one-sided message presented by a team of agents. However, the same result was not found in Experiment 2 in which all four systems were equally ranked (though there is a trend in Experiment 2 for two-sided messages and indirect communication to be less boring). As explained above, this may be due to problems with its design.

A limitation of our experiments is that they use an indirect form of self-reports where the participants judged the impact of each system on someone else's attitudes (done to ensure the arguments were relevant). Another limitation is the demographic of the participants: predominantly male in Experiment 1 and predominantly female in Experiment 2. Perhaps this has also contributed to the conflicting results.

Further experiments (with more participants) will be carried out to overcome the drawbacks of those presented in this paper, to clarify some conflicting results found and to investigate whether the trends suggested in Experiment 1 are significant. We

also plan to implement prototype systems with which the participants can actually interact so that we can measure direct effects of the systems on the participants themselves. In addition, we will explore different ways of using multiple agents, e.g. with agents playing different roles, like doctor and fitness instructor as in [13].

References

1. Andre, E., Rist, T., van Mulken, S., Klesen, M., Baldes, S.: The automated design of believable dialogues for animated presentation teams. In J. Cassell, S. Prevost, J. Sullivan, and E. Churchill: Embodied Conversational Agents. The MIT Press (2000) 220-255
2. Aronson, E.: The social animal, Worth (2004)
3. Baylor, A. L., Rosenberg-Kima, R. B., Plant, E. A.: Interface Agents as Social Models: The Impact of Appearance on Females' Attitude Toward Engineering. CHI (2006)
4. Bickmore, T.: Methodological review: Health dialog systems for patients and consumers, Journal of Biomedical Informatics Vol 39 Issue 5 (2006) 556-571
5. Cialdini, R.: Influence: science and practice. Scott, Foresman (1988)
6. Craig, S. D., Gholson, B., Garzon, M. H., Hu, X., Marks, W., Wiemer-Hastings, P., Lu, Z.: Autor tutor and otto tudor. AIED-Workshop on Animated and Personified Pedagogical Agents. Le Mans, France (1999) 25-30
7. Dehn, D. M., van Mulken, S.: The impact of animated interface agents: a review of empirical research. International Journal of Human-Computer Studies Vol 52 (2000) 1-22
8. Fishbein, M., Ajzen, I.: Belief, attitude, intention, and behavior. Addison-Wesley, Reading, MA (1975)
9. Fogg, B.J., Marshall, J., Kameda, T., Solomon, J., Rangnekar, A., Boyd, J., and Brown, B.: Web Credibility Research: A Method for Online Experiments and Early Study Results. CHI 2001 (2001) 295-296.
10. Fogg, B.J., Soohoo, C., Danielson, D., Marable, L., Stanford, J., Tauber, E.: How do users evaluate the credibility of websites?: A study with over 2500 participants. Designing for user experiences. (2003) 1-15.
11. Miller, G.R.: On being persuaded: Some basic distinctions. In M.E. Roloff & G.R. Miller (Eds.), Persuasion: New directions in theory and research. Beverly Hills, CA: Sage (1980) 11-28
12. Nass, C., Reeves, B., Leshner G.: Technology and roles: A tale of two TVs. Journal of Communication Vol 46 No 2. (1996) 121-128
13. Nguyen, H., and Masthoff, J.: Is it me or is it what I say? Source image and persuasion. Submitted to conference.
14. O'Keefe, J.D.: Persuasion: theory and research. Newbury Park, CA: Sage (1990)
15. Pew: Internet health resources, Washington, DC, Pew Internet & American Life Project (2003)
16. Riegelsberger, J., Sasse, M.A., McCarthy, J.D.: Shiny happy people building trust? Photos on e-commerce websites and consumer trust. Proceedings of CHI 2003 (2003) 121-128
17. Stiff, J.B., Mongeau, P.A.: Persuasive Communication, 2nd Edition, Guilford Press (2002)
18. de Vries, H., Brug, J.: Computer-tailored interventions motivating people to adopt health promoting behaviours: Introduction to a new approach, Patient Educ Couns Vol 36 (1999) 99-105
19. de Vries, P.: Social presence as a conduit to the social dimensions of online trust. Persuasive Technology Conference (2006) 55-59