What Makes Social Feedback from a Robot Work? Disentangling the Effect of Speech, Physical Appearance and Evaluation

Suzanne Vossen, Jaap Ham, and Cees Midden

Eindhoven University of Technology, Human Technology Interaction, P.O. Box 513, 5600 MB Eindhoven, The Netherlands
{S.H.Vossen,J.R.C.Ham,C.J.H.Midden}@tue.nl

Abstract. Previous research showed that energy consumption feedback of a social nature resulted in less energy consumption than factual energy consumption feedback. However, it was not clear which elements of social feedback (i.e. evaluation of behavior, the use of speech or the social appearance of the feedback source) caused this higher persuasiveness. In a first experiment we studied the role of evaluation by comparing the energy consumption of participants who received factual, evaluative or social feedback while using a virtual washing machine. The results suggested that social evaluative feedback resulted in lower energy consumption than both factual and evaluative feedback. In the second experiment we examined the role of speech and physical appearance in enhancing the persuasiveness of evaluative feedback. Overall, the current research suggests that the addition of only one social cue is sufficient to enhance the persuasiveness of evaluative feedback, while combining both cues will not further enhance persuasiveness.

Keywords: energy conservation, social feedback, social cues, evaluation.

1 Introduction

In the past decades, researchers have tried to find the best way to persuade people to consume less energy in their households (for an overview, see [1]). One strategy is to provide consumers with feedback about energy consumption. Several researchers concluded that especially immediate, device-specific feedback is effective in decreasing energy consumption (see e.g. [2]). Increasingly, smart meters and energy consumption displays have been deployed to make this kind of feedback possible. These technological systems can be used to persuade people in an interactive way; they function as *persuasive technology.*

Previous research suggests that the persuasiveness of such technology can be increased by adding social cues to the interface. Midden and Ham [3] provided evidence for this assumption in the field of energy consumption feedback. In a lab experiment, they compared the persuasiveness of interactive, device specific feedback from a traditional energy meter to the persuasiveness of interactive, device specific feedback from a social robot. Feedback from the robot resulted in significantly lower

T. Ploug, P. Hasle, H. Oinas-Kukkonen (Eds.): PERSUASIVE 2010, LNCS 6137, pp. 52–57, 2010.

energy consumption than feedback from the energy meter. Thereby, this research provides evidence that social feedback is more effective than factual feedback.

However, Midden and Ham [3] combined several cues to create the social feedback condition (i.e. the use of speech, the humanlike appearance of the agent and the evaluation of behavior). Although in itself it is fascinating to see that these cues enhance the persuasiveness of a feedback device, we argue that it would be even more interesting to have the relative contribution of each of the separate cues disentangled. Not only does this provide insight into the (combinations of) cues that enhance the feedback effectiveness, it also indicates which cues may be essential in evoking a social reaction to computers in general.

Therefore, the goal of the present research was to disentangle the effects of speech, humanlike appearance and evaluation. In a first experiment, we studied the effect of evaluation on feedback effectiveness by comparing the persuasiveness of factual feedback, evaluative feedback and social feedback. In a second experiment we disentangled the effects of speech and humanlike appearance on feedback effectiveness by comparing feedback including no cues, one cue, or both cues.

2 Experiment 1: Evaluation

Previously, researchers focused on the effect of *social* evaluative feedback only: evaluation originating from a social feedback source (e.g. [4, 5, 6]). The question remains whether evaluation as such is equally able to enhance the persuasiveness of feedback or that it is essential that it is *originating from another social entity* (being either a human or a social robot). Human-human interaction does not give the opportunity to answer this question, because evaluation by another person is social by nature. However, computers can be easily used to give purely evaluative feedback, for example by means of displaying colors such as red and green.

We argue though, that evaluation as such is not sufficient to evoke a social reaction to a computer because the absence of social cues may suppress the feeling that the evaluation has social consequences or is socially relevant. Therefore, we expected that evaluative feedback would not result in lower energy consumption than factual feedback. Furthermore, we expected that social evaluation would result in lower energy consumption than both evaluative and factual feedback.

2.1 Method

Participants and Design. A 3 (feedback type: social feedback vs. evaluative feedback vs. factual feedback) x 10 (washing trial one to ten) mixed design was used. Sixty students (52 men and 8 women) were randomly allocated to the three experimental groups. Participants' ages ranged from 17 through 25 (M = 19.02, SD = 1.78).

Materials and Procedure. Upon arrival, the participants were seated individually in front of a laptop computer. For participants in the embodied agent condition only, an iCat robot (see [3, 6]) was positioned on the participants' desk.

Participants were asked to complete several washing trials on a simulated washing machine panel (see [3, 6]), that was presented on the screen. They were given two goals: clean their clothes and save energy. Participants had to make a trade-off

between these goals. They were informed that during the trials they would receive feedback on how much energy the current washing program would consume. The expression of the feedback depended on the experimental condition of the participant.

The factual feedback was given by means of an energy meter that was added to the washing machine panel. It displayed the amount of consumed energy by means of a bar that increased in length when the energy consumption increased. In this condition, the behavior of the participant was not evaluated: no reference point was displayed.

Participants who received evaluative feedback were told that the computer would evaluate their energy consumption and give feedback by changing the background color of the computer screen (from green to red and vice versa).

Finally, participants who received social feedback were instructed that the iCat would tell them whether their energy consumption was good or bad while it displayed emotional expressions consistent with the given feedback.

Each participant completed a practice trial and ten washes, which were identical for all of the participants. For each trial, participants were instructed to complete a specific type of washing task (e.g., "wash four very dirty jeans").

After the participants completed all 10 washing trials, they answered several demographical questions, were debriefed and thanked for their participation.

2.2 Results

The data were analyzed by means of One-Way ANOVA in SPSS. The mean energy consumption in kWh on the ten trials served as the dependent variable.

As expected, we found a main effect of feedback type, $F(2,57) = 7.246$, $p < 0.01$. We examined the differences between the feedback types by means of linear contrasts. First, participants who received evaluative feedback ($M = .80$, $SD = .18$) did not consume less energy than participants who received factual feedback ($M = .84$, $SD = .19$), $t(57) < 1$. Furthermore, participants who received social feedback ($M = .64$, $SD = .17$) used less energy than participants who received either factual feedback, $t(57) = 3.610$, $p < .01$ or evaluative feedback, $t(57) = 2.852$, $p < .05$.

2.3 Discussion

This first study showed that evaluation as such did not result in lower energy consumption than factual feedback. This suggests that for evaluation to be effective, at least some kind of social cue needs to be present. However, it is not clear yet which social cues should be present, because in the social feedback condition two cues were combined: the physical appearance and the use of speech. We conducted a second experiment to examine the individual and combined contribution of these social cues.

3 Experiment 2: Physical Appearance and Speech

Both the use of speech and a humanlike physical appearance have been identified as social cues that are able to evoke social responses to computers [7]. Therefore, we expected that the addition of speech as well as a humanlike appearance would independently enhance the effectiveness of evaluative feedback. A more interesting

question was whether the effectiveness of social feedback was caused in particular by either speech or body or by the combination of these two.

Because speech is considered to carry strong social cues [8], we expected that in general the use of speech can enhance the effectiveness of evaluative feedback. This effect of speech probably is strongest when no other social cues are present.

Furthermore, we suggested that in case social presence is conveyed by means of speech, the physical presence of an agent does not contribute to the persuasiveness of the agent [6]. Consequently, we expected that if speech is used to give feedback, appearance would not be able to boost the persuasiveness. However, if no other social cues are present, a humanlike appearance may enhance the feedback effectiveness.

To address these issues, we studied the influence of energy consumption feedback from an embodied agent versus a computer, expressed either verbally or visually.

3.1 Method

Participants and Design. A 2 (feedback source: embodied agent vs. computer) x 2 (feedback presentation mode: verbally - speech vs. visually) x 10 (washing trial one to ten) mixed design was used. Eighty students (63 men and 17 women) were randomly allocated to the four experimental groups. Participants' ages ranged from 17 through 24 years old (M = 19.58, SD = 1.45).

Materials and Procedure. The procedure was the same as the procedure described in experiment 1, except for the way the feedback was presented.

The *spoken* feedback from the iCat was the same as the social evaluative feedback, described in paragraph 2.1.2, except for the absence of positive feedback. To generate the spoken feedback from the computer, the speech files from the iCat were played by means of the speakers of the laptop, so no agent was present.

The *visual* feedback from the computer was practically the same as the non-social evaluative feedback described in paragraph 2.1.2. However, we did not change the background color of the computer screen to give feedback. Instead, we positioned a colored LED-light behind the screen. The light was clearly visible because the color of the light was diffused on the white wall behind the screen.

Finally, participants who received feedback from the iCat visually, were instructed that the iCat would give them feedback about their energy consumption by means of changing its facial expressions and changing the color of the light. In this case the light was positioned behind the iCat. No speech was used.

3.2 Results

The data were analyzed by means of repeated measures ANOVA in SPSS. The energy consumption in kWh on each of the ten trials served as the dependent variable.

The results indicated that participants who received spoken feedback (M = .75, SD = .18) consumed less energy than participants who received color feedback (M = .84, SD = .19), $F(1,76) = 5.502$, $p < 0.05$. Besides, participants who received feedback from the social agent (M = .75, SD = .18) consumed less energy than participants who received feedback from the computer (M = .83, SD = .19), $F(1,76) = 3.81$, $p < 0.10$.

Both main effects were qualified by a significant source by presentation mode interaction, $F(1,76) = 11.802$, $p < 0.01$. Simple effect tests showed that for spoken

feedback, it did not matter whether the feedback was given by the iCat or by the computer, $F(1,76) = 1.101$, *n.s.* However, in case of feedback by means of color, the source did matter. In this case participants who received feedback from the iCat ($M = .73$, $SD = .16$) consumed less energy than participants who received feedback from the computer, ($M = .94$, $SD = .16$), $F(1,76) = 14.510$, $p < 0.00$.

Furthermore, we found that only in case of computer feedback, presentation mode did make a difference, $F(1,76) = 16.710$, $p < 0.00$. More specifically, participants who received feedback from a computer that used speech ($M = .72$, $SD = .16$) consumed less energy than participants who received feedback from a computer that used color ($M = .94$, $SD = .16$). For iCat feedback, the energy consumption did not depend on the presentation mode, $F(1,76) <1$. This indicates that only when no other social cues are present, speech is more effective than color. If a social source is present, no difference exists between speech and color. It should also be noted that feedback which combined the two social cues (the social agent and the use of speech), did not lead to lower energy consumption than feedback in which only one of these cues was present.

4 General Discussion

In two experiments we studied the combined and individual contribution of evaluation, speech and physical appearance to the effectiveness of social feedback. The first study focused on the possibility of evaluation by means of color to enhance the effectiveness of factual feedback. Although the colors red and green are generally associated with negative and positive evaluation, a computer that used these colors to give feedback was not more effective in lowering energy consumption than a computer that provided feedback by means of an energy meter. This suggests that the evaluation by means of color was not experienced as social evaluation. Instead the colors may have functioned as a tool indicating the level of the energy consumption, comparable to the way the energy meter functions. In contrast, the evaluation that was expressed verbally by the social iCat resulted in lower energy consumption than factual feedback and evaluation by means of color.

Thus, based on the results of the first experiment, we concluded that to increase persuasiveness, evaluative feedback needs to be accompanied by one or more social cues, being either the use of speech, the presence of a social robot, or both. In the second study we compared the effects of these cues on energy consumption.

We found that feedback given verbally resulted in lower energy consumption than feedback by means of color. However, a closer look at the interaction effect showed that only computer feedback benefited from the addition of speech. The effectiveness of feedback from the iCat was not increased when speech instead of color was used.

Regarding the appearance of the feedback source, a similar effect was found. In the absence of speech, participants who received feedback from the iCat consumed less energy than participants who interacted with the computer. However, in case of spoken feedback, the presence of the iCat did not enhance the feedback effectiveness.

All in all, this second study has shown that the addition of one social cue (either speech or the presence of a social robot) was sufficient to increase the persuasiveness of evaluative feedback. Our results do not suggest a need to combine them.

Furthermore, other, more subtle, social cues may also be able to evoke this kind of social reactions. In the current research our robot in itself looked social and by moving its head, eyes, and lips, the social character was amplified. Future research may focus on the effects of more unobtrusive physical cues (such as the mere presence of a motionless robot, or the effect of eyes presented on a computer screen) on persuasiveness. Even the addition of this kind of subtle cues to technological devices may be able to enhance their persuasiveness, in our case resulting in more energy saving. We made a step in this direction by demonstrating the independent influence of speech and appearance on the persuasiveness of energy feedback devices.

References

1. Abrahamse, W., Steg, L., Vlek, C., Rothengatter, T.: A review of intervention studies aimed at household energy conservation. Journal of Environmental Psychology 25, 273–291 (2005)
2. McCalley, L., Midden, C.: Energy conservation through product-integrated feedback: The roles of goal-setting and social orientation. Journal of Economic Psychology 23, 589–603 (2002)
3. Midden, C., Ham, J.: The persuasive effects of positive and negative social feedback from an embodied agent on energy conservation behavior. Paper presented at AISB 2008, Aberdeen, Scotland (2008)
4. Delin, C., Baumeister, R.: Praise: more than just social reinforcement. Journal for the Theory of Social Behaviour 24, 219–241 (1994)
5. Gaines, L., Duvall, J., Webster, J., Smith, R.: Feeling good after praise for a successful performance: the importance of social comparison information. Self and Identity 4, 373–389 (2005)
6. Vossen, S., Ham, J., Midden, C.: Social influence of a persuasive agent: The role of agent embodiment and evaluative feedback. In: Proceedings of the 4th International Conference on Persuasive Technology, Claremont, California, United States, April 26-29 (2009)
7. Fogg, B.J.: Persuasive technology: Using computers to change what we think and do. Morgan Kaufmann, San Francisco (2003)
8. Mayer, R.E., Sobko, K., Mautone, P.D.: Social Cues in multimedia learning: Role of speaker's voice. Journal of Educational Psychology 95, 419–425 (2003)