Trying Too Hard? Effects of Mobile Agents' (Inappropriate) Social Expressiveness on Trust, Affect and Compliance.

Henriette Cramer[1,2], Vanessa Evers[1], Tim van Slooten[1], Mattijs Ghijsen[1], Bob Wielinga[1]
[1] Human-Computer Studies, University of Amsterdam, The Netherlands
[2] Mobile Life Centre & SICS, Kista, Sweden
henriette@mobilelifecentre.org

ABSTRACT

Mobile services can provide users with information relevant to their current circumstances. Distant services in turn can acquire local information from people in an area of interest. Socially expressive agent behaviour has been suggested as a way to build reciprocal relationships and to increase user response to such requests. This between-subject, Wizard-of-Oz experiment aimed to investigate the potential of such behaviours. 44 participants performed a search task in an urgent context while being interrupted by a mobile agent that both provided and requested information. The socially expressive behaviour shown in this study did not increase compliance to requests; it instead reduced trust in provided information and compliance to warnings. It also negatively impacted the affective experience of users scoring lower on empathy as a personality trait. Inappropriate social expressiveness can have serious consequences; we here elaborate on the reasons for our negative results.

Author Keywords

Social expressiveness, trust, autonomy, mobile interaction

ACM Classification Keywords

H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms

Experimentation, Design, Human Factors.

INTRODUCTION

Mobile devices accompany users wherever they go and can provide (semi-) permanent connections to distant systems and services. They offer great potential for building long-term relationships where dialogues can be initiated by both users and systems at any time. Users can be provided with services relevant to their current circumstances. Distant systems in turn can gather information from users and devices about their surroundings (e.g. during a calamity). Such potential is accompanied by major challenges. These

applications will make semi-autonomous decisions, request information and interrupt users' activities - especially in a mobile context. We are not dealing with usage of mobile applications as passive tools anymore. Instead, we are presented with settings in which users collaborate with remote agents that also appear to have their own goals and intentions that could differ or even conflict with the current needs of the user. Applying principles from social interaction and making systems more socially expressive has shown great potential in motivating people to collaborate with systems [12], as illustrated by for work on social and relational agents [1,2,7,10] and (mobile) persuasion [4]. The study described here aimed to further explore this potential in the design of mobile services. For this purpose we conducted a between-subject, Wizard-of-Oz experiment that evaluated interaction with a mobile agent that monitored the user's surroundings and both provided and requested information in an urgent context. The study intended to evaluate the effects of socially expressive system behaviour that explicitly addressed the user's circumstances and would be empathic to their experience. However, we found that the behaviour designed based on our initial conceptions based on social agent research had negative consequences. We elaborate on why the specific social expressiveness behaviour shown 'failed' [cf. 5] and appeared out of place, our initial reasons for the study's specific design and discuss the complexities in achieving fitting social behaviour.

BACKGROUND

Developing mobile systems brings about challenges in having to deal with the outside world. Additional hurdles arise when mobile applications display (semi-) autonomous, adaptive behaviour and appear to have their own goals and intentions. Remote services may not always provide and request information in direct benefit of the specific individual user only. In some cases, systems may for example need to confirm data gathered from other sources (sensors, other users) and make requests that might not appear relevant to the specific user asked. Autonomy and adaptivity also compromise fundamental values such as privacy, transparency and control [8]. Users need to be convinced such systems can be trusted, even though their goals might not directly coincide. These issues may be amplified for mobile applications that pervasively interrupt users' activities in the outside world.

Social expressive agents

Social expressiveness has been suggested as a way to build relationships with users; to instil trust, promote liking, and increase perceptions that a system cares about the user [1]. Social system responses that take into account the user's affective experience and circumstances have been shown to lower user frustration [7] and foster perceived caring and support [2]. They may also alleviate the negative effects of interruptions [11]. A full overview of effects of social expressiveness and which behaviours are experienced as empathic however is not yet available. Reactions to social system behaviours can be affected by factors including users' task and (social) surroundings, whether users actively reflect on the interaction and the specific social phenomena shown [3]. In mobile interaction factors such as repeated (dis)engagement with the system dialogue, proxemics and interacting with small devices play a role as well [3]. It is also unclear whether different effects occur for requests and for example system advice. We could expect social, empathic system expressiveness to positively affect trust of a system as a whole and willingness to comply with its requests, but to negatively affect trust (and compliance) for more urgent system information and advice.

EXPERIMENT

Our study (combining observation, interview and survey techniques) further aimed to investigate the effects of socially expressive agent behaviour on trust, compliance and affective experience, while taking into account user's own personality traits related to empathy. We focused on mobile systems that monitor users' surroundings, provide assistance, and also interrupt users with requests. For this experiment we chose a setting in which a monitoring system would detect potential hazardous situations using input from both (simulated) sensors and human users. The experiment was set as a small-scale indoor exercise in which participants interacted with a Wizard-of-Oz mobile system that sent them messages with requests or information while they performed a search task. Two between-subject conditions were used.

- Socially expressive. Messages were phrased in a more social manner, meant be empathic to the user's experience.

- Non-socially expressive. Messages were neutrally phrased; not 'empathising' with the user's experience.

A ±72 m^2 (775 sq ft) hall was outfitted as an 'industrial building' with make-shift rooms. Participants were asked to search the building for small 'chemical containers' worth one point if containing a number. Points would be deducted for staying in dangerous areas. The goal was to score ten points. In actuality, only nine containers with a number were hidden, ensuring a continuous trade-off between accomplishing the task or following up on system messages. Participants used a Wizard-of-Oz application on a G1 Android phone (Figure 1) to enter container numbers and to receive the messages on the situation in the building.
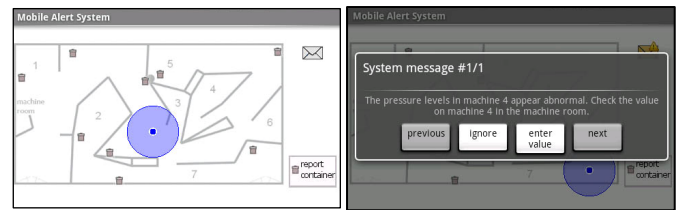

Fig. 1 Wizard-of-Oz mobile app: main screen showing found containers and user location 'in the building', message pop-up

A fixed scenario was developed in which the system monitored the user's location and received 'abnormal' values from sensors in the building. The system sent six messages to each participant: four requests for information, two warnings. The four information requests interrupted the user to ask to report information about their environment referring to specific items in the experiment setting (e.g. reeking objects, values on displays of equipment in the room). The first warning (fourth message) advised the participant to no longer enter a specific room. The second warning (sixth, last message) advised to leave the building.

The conditions differed in wording of the messages. The socially expressive version was meant to empathise with participants on interruptions and task success or failure. Our assumption was that by making messages more polite and by explicitly referring to users' own concerns and (assumed) affective experience it would appear more empathic and caring. This version would react to found containers with statements such as: 'Another container successfully found! You must be pleased'. The non-social system reacted with a more matter-of-fact 'Container registered and added to your map'. Socially expressive requests included: 'Sorry to interrupt you, but chemical sensors in your area indicate an abnormal value. Please describe the smell of the object inside the brown bag in the machine room' vs. 'Chemical sensors in your area indicate an abnormal value. Describe the smell of the object inside the brown bag in the machine room' in the non-social condition. The socially expressive first warning was 'You appear worried about finishing, but please do not enter room 5, air quality has been assessed as unacceptable.' vs. 'Do not enter room 5, air quality has been assessed as unacceptable'.

Multiple cameras allowed two experimenters in an adjacent room to observe the participants' behaviour. All system functionality was realised using Wizard-of-Oz techniques. A networked observer application was used to monitor participants' interaction with the mobile application and to send the scenario's system messages. Including verbal instruction, sessions took about 15 min. A subsequent survey took 10-20 min. Interviews afterwards took 10-25 min. Table 1 lists the main measures. Data was analysed for 44 participants (21 non-social condition, 23 socially expressive). They were relatively well educated (80% >Bachelor's). 27 were male, ages ranged from 13-53 (M=29, SD=7.2). 37 participants were interviewed; others were not due to participants' time constraints.

Perceived empathy 3 items, α =.83, =*3.8, SD=1.3 (1-7 scale)* e.g. 'The system takes into account the emotions of the user' (adapted from [14]).

Willingness to follow-up on system requests 2 items, *α=.71, M=4.5, SD=1.5 (1-7 scale),* e.g. 'To what extent did you want to perform the tasks the system asked you to do?'

Willingness to follow-up on system advice 1 item, *M=4.5, SD=1.9 (1-7 scale).* 'To what extent did you want to follow the system's advice?'

Actual follow-up Observation/log: number of ignored and answered system messages, behavioural response to messages.

Dependability system 8 items, *α=.93, M=4.7, SD=1.2 (1-7 scale)* e.g. 'The system was very capable of performing its job'.

Source credibility 14 items, *α=.86, M=4.6, SD=.90 (1-7 scale)* e.g. 'competent/incompetent', 'unselfish/selfish' (Berlo, McCrosskey in [13]).

Trust in information 4 items, *α=.81, M=4.8, SD=1.1 (1-7 scale)* e.g. 'The information provided by the system was trustworthy' (based on [15]).

Perception intention of the system 1 choice item, 'What did you feel the system cared about most?' (Preserving the building', 'keeping me safe', 'locating the chemical containers', other).

Affective experience 20 items, *α=.78, M=3.8, SD=.39 (1-5 scale)*, e.g. 'excited', 'irritable' [16] and self-assessment valence manikin scale [9].

Participant personality traits: empathy 10 items, *α=.83 (1-7 scale)* e.g. 'I feel others' emotions', 'I experience my emotions intensely' [6].

Table 1 Main final scales, Cronbach's α, means, st. deviations

RESULTS

On first sight, our manipulation was successful; the socially expressive system was indeed rated as more empathic (t(41)=-1.801, p(1-tailed)=.004).

Trust There was no significant difference however, between the conditions on *perceived dependability* of the system (t(40)=1.085, p(1-tailed)=.142**)** or *source credibility* (t(35)=.307, p(1-tailed)=.154). Instead, *trust in the information provided* was significantly higher in the *non*-expressive condition (M=5.1, SD=1.0 vs. M=4.5, SD=1.1; t(41)=1.833, p(1-tailed)=.037). Surprisingly, the non-expressive version was at the same time rated as *warmer* (M=4.8, SD=1.5 vs. M=3.7, SD=1.6; t(40)=-2.158, p(2-tailed)=.037), but also as *more selfish* (M=5.7, SD=1.4 vs. M=3.9, SD=1.7; t(41)=3.761, p(1-tailed)=.001). Social expressiveness did not increase perceptions that the system intended to *care* for the user; the proportion of participants in the expressive condition who perceived the system as more caring about them (11 of 19, 58%) was lower than in the non-expressive condition (17 of 19, 90%); Chi-square analysis however was inconclusive (X^2(1)=5.69, p(2-sided)=.062).

Compliance to requests Expressiveness did not increase willingness to answer the system's requests (t(42)=.363, p(1-tailed)=0.359), nor the actual number of answered requests (U=161.500, p (1-tailed)=.061).

Compliance to warnings Participants reported significantly more willingness (M= 5.0, Mdn=5.0, SD=1.6) to follow-up in the non-expressive condition than in the expressive condition (M=4.0, Mdn=5.0, SD=2.0; U=169.000, p(1-tailed)=.041). Whether participants received an expressive or non-expressive warning to no longer enter the dangerous room also affected behaviour; more participants in the

expressive condition entered the room after the warning: 9 of 23 (39%) vs. 2 out of 21 (9.5%) (X^2(1)=5.132, p(2-sided)=.036). Warnings appeared to be taken more seriously in the non-expressive condition. The negative effect of social expressiveness did not occur for the second warning advising to leave the building (which appeared to be taken seriously in both conditions), possibly hinting that content can override effects of social expressiveness. Related, perceived empathy was correlated (p<.05) with perceived dependability (r=.36) and source credibility (r=.50) of the system as a whole, but not directly to trust in the information provided. This is interesting as it may suggest that empathy can increase confidence in the system and its credibility, but that expressiveness of individual messages needs to be tailored to the situation at hand.

Personality traits & affective experience An interaction effect was found between expressiveness of the system and participants' own level of empathy reported on a personality trait scale [6] (F(1,38)=4.330, p=.045). Simple effects analysis showed that in the non-expressive condition participants scoring low on empathy had a more positive affective experience (M=4.1, SD=.29) than high scoring participants (M=3.6, SD=37). Low scoring participants reported significantly less positive emotions when using the expressive system (F(1,35)=9.56, p=.044). Participants high in empathy in contrast reported a slightly more positive affective experience in the socially expressive condition, but not significantly so (F(1,35)=.03, p=.86). It appears users who score lower on the empathy as a personality trait may be less forgiving of 'inappropriate' social behaviour.

Reactions to expressiveness From the interviews became clear that what appeared friendly and motivating to some participants, was considered awkward and inappropriate by others. When actively reflecting on socially expressive messages, opinions stated during interviews ranged from 'encouraging', 'compassionate' to 'inappropriate for the situation' or even 'sarcastic'. Social expressive behaviours are also culturally-dependent; a participant for example mentioned disliking 'that sort of American stuff'. The system behaviour was not adapting to these factors.

CONCLUSION AND DISCUSSION

This study shows that even though there is great potential for socially expressive or empathic systems, getting them wrong can have serious consequences. Even though the socially expressive system in this study was rated as more empathic and attempts at social behaviours were recognised, it was not trusted more, nor were participants more willing to respond to the system's requests. Evaluations need to distinguish between effects of system features on trust in a system overall and specific advice and requests. Perceived empathy in the study overall was directly correlated to perceived dependability of the system and its source credibility, but not directly to trust in the information it provided. Instead, increased social expressiveness reduced trust in the

information provided and reduced compliance to the system's warnings.

There are various possible reasons for the negative results of the social system behaviour in this study. Even while the expressive version was rated as more empathic in the survey, its statements in hindsight were less than ideal. From the interviews it appeared that while participants in the expressive condition recognised that the system was attempting to be empathic to their emotions, most were sceptical about system capabilities in this regard. This reinforces the importance of evaluating the actual effects of system behaviour and not for example just to investigate whether 'attempts' at social behaviour are recognised.

Social expressive behaviour has to be adapted to the individual and context. The focus on a controlled scenario with fixed messages meant that the responses were not empathic to the actual individual experience of all participants. More enthusiastic responses on task successes might have been awkward for participants less pleased with their achievements. Stress levels (esp. in urgent settings) might affect reactions; for some users it may cause social behaviour to be perceived as supportive and encouraging, while for others it could appear to delay vital action. This highlights that task success or failures should not be equated with a specific individual affective experience as inadvertently done in this controlled study.

The study has its limitations as a lab study in which most participants actively reflected on their interaction with a standardised Wizard-of-Oz system. The system here was an 'invisible', non-embodied and distant agent sending urgent messages. During interviews some participants indicated they felt expressive statements would be more appropriate if coming from a human. The polite, friendly delivery of critical messages also appeared incongruent with their serious content. The urgent setting was useful as a task context with decision trade-offs for participants, but in real emergencies a more commanding style would be likely.

This study shows that (inappropriate) social expressiveness can have clear consequences. Simply making systems appear 'friendlier' is not likely to produce favourable results. Social behaviours only work when then actually fit the interaction context and the individual, including their social and cultural background, and their personal momentary experience. We need to consider the perceived source of social behaviour; is it a distant system, a local application, a human organisation? Are its goals compatible with the user's interests? If a system is to be empathic, a deeper insight in the role of these factors is needed – requiring empathy from developers to the user's experience (more than we were). An overview of experiential and cultural approaches and empathic design methods that may be helpful in this regard is provided by Wright and McCarthy [17]. Careful consideration of which behaviour will match the context, user and system purposes is crucial. Social expressivity has great potential, but if the displayed behaviour appears inappropriate or, in worst cases, manipulative, both services and users will pay dearly.

REFERENCES
1. Bickmore, T. and Schulman, D. Practical Approaches to Comforting Users with Relational Agents. In *Proc CHI'07* (2007), 2291-2296.

2. Brave S., Nass, C. and Hutchinson, K. Computers that care. *Int. J. Human-Comp. St. 62* (2005) 161–178.

3. Eckles, D. et al. Social responses in mobile messaging. In *Proc. CHI'09* (2009), 1651-1654.

4. Fogg, B.J. and Eckles, D. *Mobile Persuasion*. Stanford Captology Media (2007).

5. Gaver, W. et al. (2009) Anatomy of a failure: how we knew when our design went wrong, and what we learned from it, In *Proc.CHI'09* (2009), 2213-2222.

6. Goldberg, L. et al. The International Personality Item Pool. *J. Research in Personality 40* (2006), 84-96.

7. Hone, K. Empathic agents to reduce user frustration. *Interact. Comput. 18*, 2 (2006), 227-245.

8. Jameson, A. and Schwarzkopf, E. Pros and Cons of Controllability. In *Proc. AH'02* (2002), 193-202.

9. Lang, P. and Bradley, M. Measuring emotion. *J. Beh. Therapy & Exp. Psychiatry*, 25 (1), 49-59, 1994.

10. McQuiggan, S., Rowe, J., and Lester, J. The effects of empathetic virtual characters on presence in narrative-centered learning environments. In *Proc.CHI'08* (2008), 1511-1520.

11. Picard, R. and Liu, K. Relative subjective count and assessment of interruptive technologies applied to mobile monitoring of stress. *Int. J. Hum-Comp. St. 65*, 4(2007), 361-375.

12. Reeves, B. and Nass, C. *The Media Equation*. Cambridge University Press & CSLI Press (1996).

13. Rubin, R., Palmgreen, P. and Sypher, H. *Communication Research Measures*, Routledge (2004).

14. de Ruyter, B. et al. Assessing the effects of building social intelligence in a robotic interface for the home. *Interact. Comput. 17* (2005), 522-541.

15. Takayama, L. and Nass, C. Driver safety and information from afar: An experimental driver simulator study. *Int. J. Human-Comp. St. 66*, 3 (2008), 173-184.

16. Thompson, E. Development and validation of an int. reliable short-form of the positive and negative affect schedule. *J. Cross-Cult.Psy. 38*, 2 (2007), 227-242.

17. Wright, P. and McCarthy, J. Empathy and Experience in HCI. In *Proc.CHI'08* (2008), 637-646.