Do People Trust Their Eyes More Than Ears?
Media Bias in Detecting Cues of Expertise

Jens Riegelsberger, M. Angela Sasse, John D. McCarthy
Department of Computer Science
University College London
Gower Street, London WC1E 6BT, UK
{jriegels, a.sasse, j.mccarthy@cs.ucl.ac.uk}

ABSTRACT

Enabling users to identify trustworthy actors is a key design concern in online systems and expertise is a core dimension of trustworthiness. In this paper, we investigate (1) users' ability to identify *expertise* in advice and (2) effects of *media bias* in different representations. In a laboratory study, we presented 160 participants with two advisors – one represented by *text-only*; the other represented by one of four alternate formats: *video*, *audio*, *avatar*, or *photo+text*. Unknown to the participants, one was an *expert* (i.e. trained) and the other was a *non-expert* (i.e. untrained). We observed participants' *advice seeking* behavior under financial risk as an indicator of their trust in the advisor. For all rich media representations, participants were able to identify the expert, but we also found a tendency for seeking *video* and *audio* advice, irrespective of expertise. *Avatar* advice, in contrast, was rarely sought, but – like the other rich media representations – was seen as more enjoyable and friendly than text-only advice. In a future step we plan to analyze our data for effects on advice uptake.

Author Keywords

Trust, CMC, interpersonal cues, video, audio, avatar, photo

ACM Classification Keywords

H5.1 Multimedia Information Systems: Animation, Audio, Evaluation, Video; H5.2 User Interfaces: Evaluation, H5.3: Group and Organization Interfaces: CSCW, Evaluation.

INTRODUCTION

As technology-mediated interaction gradually replaces face-to-face (f-t-f) interaction in many areas of life, trust becomes a central concern for designers and researchers. Systems should not be designed simply to increase user trust, but to enable users to discriminate between trustworthy and less trustworthy actors [3]. To date, research investigating users' ability to discriminate mainly focused on deception (e.g. [4,6]). However, in many

everyday situations, questions of trust do not arise from the risk of willful deception, but because one is uncertain about the other's expertise [1,2,3]: an individual might mean well, but lack the expertise to be truly helpful. Investigating these issues, we focus on the perception of interpersonal cues of expertise in advice given in the rich media representations *video*, *avatar*, *audio*, *photo+text*, and – for baseline comparisons – *text-only*. These media are chosen for their practical relevance: with more bandwidth available to users, video and audio are becoming increasingly common. Avatars and animated assistants are now marketed as cost effective off-the-shelf solutions to enrich the user experience. Photos are simple additions that have long been used online with the aim to build trust.

For these media, we are investigating (1) whether they introduce a *media bias,* and users' *ability to discriminate* between *expert* and *non-expert* advice (2). Bias occurs when advice is preferred due to its media format, irrespective of its expertise. We are also investigating the influence of *risk*. After an overview on online trust research, we introduce our predictions and experimental approach. Then we present and discuss the results of the study and we close with conclusions for researchers and practitioners.

BACKGROUND AND RESEARCH QUESTIONS

Trust has been defined as a willingness to be vulnerable based on positive expectations [1]. This implies that trust is required in the presence of *risk* and *uncertainty.* Relying on an online advisor can pose several risks, ranging from lower than expected entertainment (e.g. in the case of a film recommendation) to bodily harm (e.g. in the case of medical advice). Uncertainty arises from the fact that the trustor cannot directly observe the trustee's *ability* (e.g. expertise) and *motivation* (e.g. desire to deceive) [2], but needs to infer these from the available information. Interpersonal cues are an important type of signals for trustworthiness in f-t-f situations [5]. They include visual cues (e.g. appearance, facial expressions) and audio cues (e.g. pitch, modulation) [5].

If interactions are mediated, some interpersonal cues are lost. Text chat, for instance, removes all visual and audio cues. In the discussion on online trust, it is often assumed

that this reduction in cues will result in lower trust. However, there is also evidence that trust cannot be linked unequivocally to such a one-dimensional understanding of media richness. Firstly, in the presence of cues for lack of expertise (e.g. nervousness), a rich channel is unlikely to result in a high level of trust compared to one that suppresses such cues. Secondly, Walther [10] found that narrow-bandwidth channels can result in over-reliance on the few cues available, which may lead to unwarranted high levels of trust. Hence, richer representations may result in **(P1)** *positive media bias* (i.e. more trust) or they may result in **(P2)** *better discrimination* between expert and non-expert advice as they convey more information.

Video and Audio: Swerts et al. [9] in a study on interpersonal cues of uncertainty found that users' ability to discriminate was lowest for video-only, higher for audio-only and highest for video+audio; thus supporting P2. Investigating the detection of deception in video, Horn et al. [6] found that slight visual spatial degradation reduced participants' ability to discriminate; giving further support to P2. However, severe degradation of the visual channel resulted in better discrimination. They hypothesized that this effect may result from a reduced bias in the absence of recognizable visual cues. Such an effect would provide support for P1 and suggest that visual cues in particular introduce a positive bias.

Avatar: Virtual humans (avatars and animated assistants) are sometimes presented as simple means to enrich user experience and build trust. However, they can prompt mixed reactions from users [3]. In a study that varied agent implementation and expertise (albeit not the interpersonal cues given off) van Mulken et al. [7] found a strong effect of expertise on perceived trustworthiness but only a marginally positive effect for the embodied representation. Rickenberg & Reeves [8] found a positive effect of a simple animated agent on user trust (P1).

Photos: Photos do not give additional cues with individual advice, but they are widely used with the aim to increase social presence and trust. Previous studies found that they can bias users' trust in websites [3] (P1).

None of the studies above induced risk to measure trust and none systematically investigated P1 and P2 across different media representations. To specifically address our predictions we contrasted expertise and media richness: we gave each participant two advisors – one in a *rich media* representation and the other *text-only*. For one group of participants *expert* advice was given by the *rich media* advisor; for the other group by the *text-only* advisor. On each question participants could ask only one advisor. Figure 1, illustrates P1 and P2 for this approach. In the hypothetical case of total bias (P1), we would expect participants to always seek rich media advice, irrespective of expertise. In the case of perfect discrimination (P2), participants would always prefer expert advice.
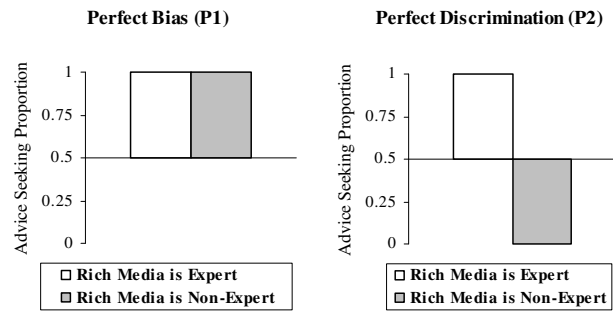

Figure 1. Illustration of predictions P1 and P2.

METHOD

Participants and Design

160 highly computer-literate participants (median age 23.75, 49% female) took part in the study, which was framed as a quiz, similar to the well-known TV show *Millionaire*. The questions used in the study had been tested for their difficulty in a pre-study with 80 participants. After two easy practice questions on which only correct advice was given, participants went through 30 assessed questions. Feedback on the correctness of participants' answers was only given at the end of the study.

The study had a 4 (*type of rich media representation*) x 2 (*rich media advisor is expert vs. rich media advisor is non-expert*) between-subjects design. Each participant was presented with a pair of advisors (Fig. 2). In all conditions one advisor was represented as *text-only* and the other in one of the four *rich media representations (video, avatar, audio, photo+text*; Fig. 2, 3). Depending on the condition, either the *text-only* or the *rich media* advisor gave *expert* advice, while the other gave *non-expert* advice. The order of the questions and answer options (A-D, Fig. 2) was randomized; the position (left, right) and names (Katy, Emma) of the advisors were counterbalanced.
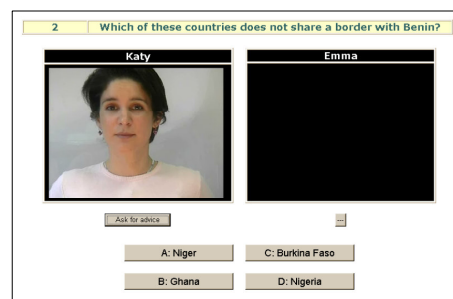

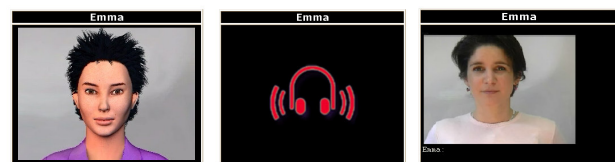**Figure 2. Experimental system (*video* and *text-only* advisor).**

Figure 3. *Avatar*, *audio*, and *photo+text* advisor.

Independent Variables

Expertise: *Non-Expert* and *expert* advice was recorded from the same individual before and after training, respectively. Hence, the *expert* and *non-expert* advisors only differed in the ratio of correct to incorrect advice and in their cues to confidence about the answers. In the interest of ecological validity, answer formats were not prescribed. Based on experience with a pilot study, we added 6 incorrect (and less confident) answers from the untrained recording to the *expert* so she did not seem artificially perfect. The proportion of correct (and confident) advice was .80 for the *expert* and .36 for the *non-expert*.

Media Representation: The media representations were created from the same video clips ranging from 1 sec. to 8 secs. in length. The original clips were used for the *video* representation. The *avatar* was created with a commercially available animation tool (V1 by DA Group) directly from the audio stream without any manual scripting of nonverbal behavior. The tool synchronized lip movements and added cues of liveliness (e.g. blinks). *Video* and *avatar* were streamed with Windows Media Encoder (350 kbps, 320x240). *Audio* was encoded with 48 kHz, 16 bit, mono. *Photo+text* included a facial photo of the advisor, otherwise it was identical to the *text-only* representation; for both, the text appeared dynamically with a delay of 107 ms per letter to ensure that all representations had equal 'playing time'.

Risk: Participants' pay depended on the number of correctly answered questions and thus on their ability to identify the expert advisor, as the quiz questions were extremely difficult. Pay ranged from the equivalent of $15 to $26. To investigate the effect of level of risk, we included a high-stakes question worth an additional $5.50.

Dependent Variables

The measure *advice seeking* was defined as the proportion of one advisor being asked out of the total number of times advice was sought by a participant. As only one advisor could be asked on each question, *expert advice seeking* = 1 – *non-expert advice seeking*. A final questionnaire elicited users' subjective assessments of the two advisors.

RESULTS

On average, participants sought advice on 26 out of 30 questions. One participant did not seek advice at all. Figure 4 shows a main effect for *expertise* (P2) on participants' *advice seeking* ($F (1, 154) = 51.56, p < .001$). There is also some indication for an effect of the type of *rich media* representation ($F (3, 154) = 2.50, p = .062$). To conduct within-subject tests for *bias* (P1, Fig.1) and *discrimination* (P2, Fig.1) in individual conditions, we investigated *rich media non-expert* advice seeking (grey bars in Fig. 4). A value < .5 would provide evidence for discrimination (see Fig. 1), a value > .5 would be a sign of bias outweighing discrimination. Figure 4 shows *non-expert avatar* and *photo+text* advice seeking significantly below .5. No such effect was present for *video* and *audio*, indicating that a media bias interferes with users' ability to discriminate.
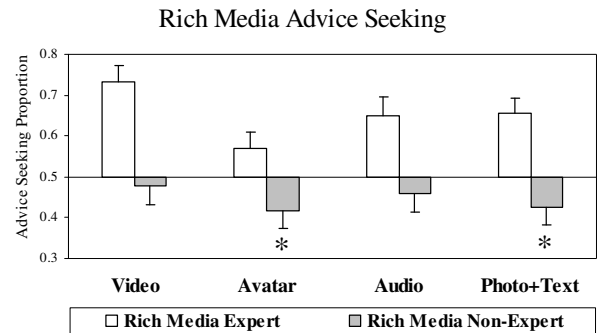

Figure 4. Seeking advice from the rich media advisor. Stars (*) indicate results for one-sided t-tests (H: *non-expert* advice *seeking* < .5; *p* < .05).

Further evidence for a media bias in *video* and *audio* is given by the finding that for these representations *rich media expert* advice was chosen more often than *text-only expert* advice (*video*: $t(38) = 3.60, p < .001$, *audio*: $t(37) = 1.69, p < .05$; both one-sided). This effect was not present for the *avatar* and *photo+text* representations. *Expert avatar* advice was less often sought than advice from the other *rich media experts* combined ($t (77) = 2.45, p < .05$).

Participants increasingly sought advice from the *expert* as they gained experience with the advisors (Fig. 5), but the increase in financial risk for the final high-stakes question resulted in an increase in seeking advice from the *rich media* advisors (McNemar: $\chi^2 (131) = 6.25, p=.012$, Fig. 6).
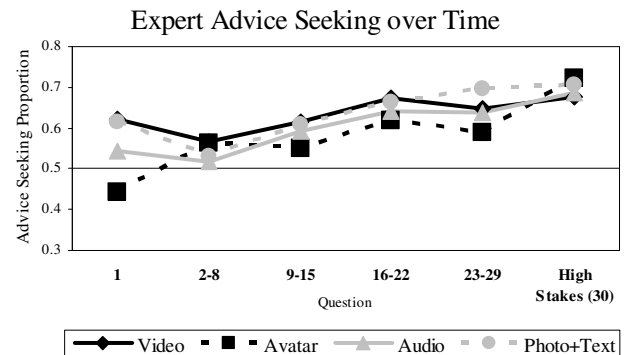
Figure 5. Expert (*rich + text-only*) advice seeking over time.
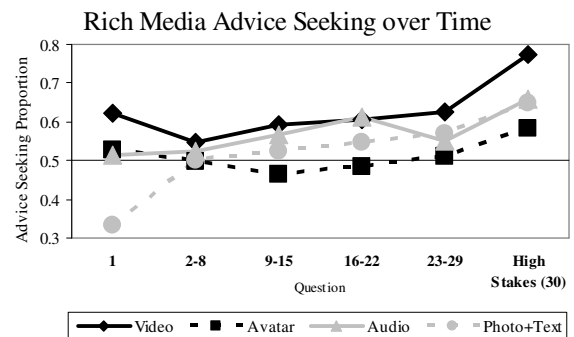
Figure 6. Rich (*expert + non-expert*) advice seeking over time.

Participants stated that they trusted the *video* advisor more than the *text-only* advisor, irrespective of *expertise* (t (39) = 2.83, $p < .01$); a finding not replicated for the other media representations. All *rich media* representations were rated as more friendly (t (159) = 7.24, $p < .001$) and enjoyable (t (159) = 6.71, p < .001) than *text-only*.

DISCUSSION

Overall, we found that participants mostly chose *expert* advice in all media representations (P2). However, there was also some indication that *video* and *audio* representations can interfere with users' ability to discriminate effectively (P1). Increased risk led to an increase in media interference.

Video: When the *non-expert* was represented in *video,* preference for choosing *video* almost matched the preference for choosing *expert* advice. Also, in the post-experimental ratings, participants stated that they trusted the *video* advisor more, irrespective of expertise. Hence, with a view to well-placed trust, video can be seen as problematic: users' preference for receiving *video* advice led them to disregard better *text-only* advice.

Audio: Similar to *video*, the preference for seeking *non-expert audio* advice almost matched the preference for *expert* advice. However, participants did *not say* they trusted *audio* more than *text-only* irrespective of expertise. This finding supports Horn et al. [6] in that visual interpersonal cues in particular appear to induce a bias.

Avatar: The *avatar* did not result in a bias; rather it was less preferred than other *rich media expert*s. This finding cannot necessarily be generalized to other avatars or animated assistants, but it indicates that using an off-the-shelf avatar to increase trust may not be advisable at this stage. Finally, we did find that the *avatar,* like all other *rich media* representations, was perceived as friendlier and more enjoyable than the *text-only* advisor.

Photo+Text: Lexical cues alone, as given in the *photo+text* representation, were sufficient for identifying the *expert*. The photo was not found to bias advice seeking, but it did result in higher ratings of friendliness and enjoyment compared to *text-only* advice.

CONCLUSIONS

We observed participants' advice seeking in a situation of limited advice and under financial risk. In all media representations, participants were able to identify *expert* advice (P2), but the data suggest that *video* and *audio* representations can interfere with users' ability to discriminate effectively (P1). One interpretation of this finding is that users chose the rich media representations because they considered them to give the best insight into the trustworthiness of a piece of advice. An analysis of our data for effects on advice uptake will clarify whether this is really the case. The relatively good performance at perceiving expertise in the *photo+text* representation suggests that sufficient information about expertise was contained in text alone.

For designers interested in high levels of trust (even at the risk of inducing media bias), *video* is the best representation, followed by *audio*. Finally, the *avatar*, and even just a simple *photo* lead to higher ratings of friendliness and enjoyment than *text-only*. So, if the design goal is engagement rather than trust, our data suggests that these representations can be effective.

With a view to methodology, our results provide further support for measuring trust by observing decision-making under risk, since we found that the level of financial risk influenced participants' behavior. Our next step will be to analyze the data for effects on participants' advice uptake.

REFERENCES
1. Corritore, C. L., Kracher, B., and Wiedenbeck, S., On-line trust: concepts, evolving themes, a model. *International Journal of Human Computer Studies*, 58, 6 (2003), 737-758.
2. Deutsch, M., Trust and suspicion, *Journal of Conflict Resolution*, 2, 3, (1958), 265-279.
3. Fogg, B. J., *Persuasive Technology*, Morgan Kaufmann, San Francisco, CA, US, 2003.
4. Hancock, J. T., Thom-Santelli, J., and Ritchie, T., Deception and Design: The Impact of Communication Technology on Lying Behavior, *Proceedings of CHI 2004*, ACM Press, New York, (2004), 129-134.
5. Hinton, P. R., *The Psychology of Interpersonal Perception*, Routledge, London, 1993.
6. Horn, D. B., Olson, J. S., and Karasik, L., The Effects of Spatial and Temporal Video Distortion on Lie Detection Performance, *CHI 2002 Extended Abstracts*, ACM Press, New York, (2002), 716-718.
7. van Mulken, S., Andre, E., and Müller, J., An empirical study on the trustworthiness of life-like interface agents. *Proceedings of HCI International '99*, 2, Lawrence Erlbaum, Mahwah, NJ, (1999), 152-156.
8. Rickenberg, R and Reeves, B. The effects of animated characters on anxiety, task performance, and evaluations of user interfaces. *Proceedings of CHI 2000*, ACM Press, New York, (2000), 49-56.
9. Swerts, M., Krahmer, E., Barkhuysen, P., and van de Laar, L., Audiovisual cues to uncertainty, *Proceedings of the ISCA Workshop on Error Handling in Spoken Dialogue Systems*, Chateau-D'Oex, Switzerland, 2003.
10. Walther, J. B. Visual Cues and Computer Mediated Communication: Don't look before you leap. *Proceedings of the Annual Meeting of the ICA,* San Francisco, CA, US, 1999.