# *Rich Media, Poor Judgement? A Study of Media Effects on Users' Trust in Expertise*

**Jens Riegelsberger, M. Angela Sasse, John D. McCarthy**

*University College London, Dept. of Computer Science*
*Gower Street, London WC1E 6BT, United Kingdom*
*Tel: +44 20 7679 0351*
*Fax: +44 20 7387 1397*
*Email: {jriegels, a.sasse, j.mccarthy}@cs.ucl.ac.uk*
*URL: www.cs.ucl.ac.uk/staff/jriegels*

**In this paper, we investigate how interpersonal cues of expertise affect trust in different media representations. Based on a review of previous research, richer representations could lead either to a *positive media bias* (P1) or *increased sensitivity for cues of expertise* (P2). 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* and *advice uptake* to infer their *sensitivity to correct advice* in a situation of financial risk. We found that most participants preferred seeking advice from the *expert*, but we also found a tendency for seeking *audio* and in particular *video* advice. Users' self-reports indicate that they believed that video in particular would give them the most detailed insight into expertise. Data for advice uptake, however, showed that all media representation, including text-only, resulted in good sensitivity to correct advice.**

**Keywords:** Trust, Expertise, Video, Avatar, Audio, Photo, CMC, CSCW

## 1 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 providers of online services (Corritore et al. 2003). In this context, many researchers investigate how to maintain or increase levels of trust. However, it is also crucial to ensure that

users are able to place trust correctly (i.e. are able to discriminate between trustworthy and less trustworthy actors). Experiencing the consequences of misplaced trust can undermine future willingness to interact with online services and technologies. To date, research investigating the correctness of trust decisions mainly focused on deceptive behaviour (e.g. Horn et al., 2002). However, in many everyday situations, questions of trust do not arise from the risk of wilful deception, but because one is uncertain about the other's expertise (Deutsch, 1958). An individual might mean well, but lack the expertise to be truly helpful. Investigating these issues, we focus on cues of expertise, a thus far under-researched constituent of trustworthiness.

Due to bandwidth constraints, online services used to be limited to providing most users with text communication or simple web pages containing photos. However, wide availability of broadband access now allows services to be delivered in richer formats, such as audio or video. In addition, avatars (animated human-like characters) now promise social presence (Short et al. 1976) at a level similar to that provided by video – albeit at a lower cost in terms of production and bandwidth. The four *rich media formats* we examine in this paper are *video, avatar, audio*, and *photo + text*. As a baseline measure for comparisons we includea *text-only* condition. We are particularly interested in how different representations affect users' *sensitivity to cues of expertise*, i.e. the degree to which they can identify correct and incorrect advice. We are investigating whether richer representations result in either a media bias (1) or increased sensitivity for cues of expertise (2). Bias occurs when advice is preferred due to its media format, irrespective of its expertise.

After an overview of online trust research (Section 2), we introduce our predictions and methodological approach (Section 3). Then we present and discuss the results of an experimental study that was conducted to test our predictions (Sections 4 and 5). We close with conclusions for researchers and practitioners (Section 6).

## 2 Background

### *2.1 Trust and Interpersonal Cues*

Trust has been defined as a willingness to be vulnerable, based on positive expectations (Corritore et al. 2003). This implies that trust is required in the presence of *risk* and *uncertainty* (Corritore et al. 2003; Giddens 1990; Deutsch 1958). 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), but needs to infer those from cues (Bacharach & Gambetta 1997). *Interpersonal cues* can play an important role in the perception of trustworthiness in f-t-f situations, because they give information about an individual's background (e.g. education, provenance), but also about intrinsic states such as sincerity and confidence (Whittaker & O'Conaill 1997; Zuckerman et al. 1981). Interpersonal cues include visual cues (e.g. appearance, facial expressions) and audio cues (para-verbal: e.g. pitch; Hinton 1993).

If interactions are mediated, some interpersonal cues are lost. Text chat, for instance, removes all visual and audio cues. Intrinsic states and personal background can only be inferred from vocabulary and phrasing. In the view of media richness models (Rice 1992), text chat is considered to result in low social presence and is thus seen as a poor channel. In the discussion on trust, it is often implicitly assumed that a poor channel will result in lower trust, as many of the interpersonal cues that are crucial for building trust are not present (Giddens 1990; Handy 1995). Visual interpersonal cues (e.g. smiles), which are suppressed by text or audio representations, have been identified as particularly powerful in evoking immediate affective responses (Winston et al., 2002).

However, there is also evidence that trust cannot be linked unequivocally to a one-dimensional model of media richness. In the presence of cues for untrustworthiness (e.g. nervousness), a rich channel is unlikely to result in a high level of trust compared to one that suppresses such cues. Walther (1999) found that narrow-bandwidth channels can also result in over-reliance on the few cues available, and thus may lead to unwarranted high levels of trust.

Two predictions regarding the effect of media richness on trust are therefore possible: richer representations may result in **(P1)** *positive media bias* (i.e. more trust) because they increase social presence or they may result in **(P2)** *better discrimination* between trustworthy and less trustworthy actors as they convey more information.

## *2.2 Evidence for Media Bias (P1) and Discrimination (P2)*

We briefly review trust research that specifically addressed video, audio, avatars and photos with a view to P1 and P2.

**Video.** In social dilemma studies, video resulted in the highest levels of cooperation when compared to audio and text-only communications (Bos et al., 2002; Olson et al., 2002), thus providing some evidence for P1. In a study on interpersonal cues of uncertainty, Swerts et al. (2004), however, 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. (2002) 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. Horn et al. (2002) 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.

**Avatars.** Virtual humans (avatars and embodied agents) are sometimes presented as simple means to enrich user experience and build trust. They can be easily produced with off-the-shelf tools from an audio stream. However, they can prompt mixed reactions from users depending on implementation, context, and user characteristics (Fogg 2003). In a study that varied agent implementation and expertise (albeit not the interpersonal cues given off) van Mulken et al. (1998) found a strong effect of expertise on perceived trustworthiness but only a marginally positive effect for the embodied representation.

**Audio.** In line with media richness models, audio-only communication in social dilemma studies resulted in levels of cooperation that were lower than those for video, but higher than those found for text-only communications (P1; Bos et al. 2002; Olson et al. 2002). Even synthetic speech was found to reduce uncooperative behaviour compared to text chat. Davis et al. (2002) attribute this finding to the social presence afforded by synthetic voice. Swerts et al. (2004) on the other hand found that audio-only allowed better discrimination than video-only, suggesting that audio cues in particular give insight into certainty (P2).

**Photos.** Photos do not give additional cues with individual advice compared to text-only representations (P2), 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 (P1; Fogg 2003).

None of the studies above induced risk to measure trust and at the same time systematically investigated P1 and P2 across different media representations. Hence, to specifically address these predictions, we designed a study that contrasted expertise and media richness. We modelled our experimental study on a user-advisor relationship, a widely used research paradigm in social psychology (Yaniv & Kleinberger 2000) and gave participant *expert* and *non-expert* advisors, one of them *text-only* and the other in rich media representations (see Section 3.1). The study was framed as a general knowledge quiz, similar to the well-known TV show '*Who Wants to Be a Millionaire?*'

# 3 Method

## *3.1 Participants and Design*

160 participants took part in the study. The median age was 23.75 (SD = 3.30) and the sample was balanced for gender (49 % female). The study had a 4 media (*type of rich media representation*) x 2 expertise (*rich media advisor is expert vs. rich media advisor is non-expert*) design, resulting in 8 between-subject conditions with 20 participants each (Table 1).

| **Advisor 1** | **Video** | **Avatar** | **Audio** | **Photo+Text** |
|---|---|---|---|---|
| **Advisor 2** | **Text-only** | **Text-only** | **Text-only** | **Text-only** |
| Advisor 1 is the expert | 20 | 20 | 20 | 20 |
| Advisor 2 is the expert | 20 | 20 | 20 | 20 |

**Table 1. 8 between-subject conditions used in the study.**

In each between-subject condition, two advisors were available (Figure 1) – one represented as *text-only,* and the other in one of the *rich media representations*. The rich media representations were *video*, *avatar*, *audio*, and *photo+text*. Depending on the factor *expertise*, 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) was randomized; the position (left, right) and names (Katy, Emma) of the advisors were counterbalanced.

Prior to starting the assessed part of the experiment, participants completed two training rounds that consisted of easy questions. For these, both advisors gave identical and correct advice. Then participants answered 29 assessed questions, followed by a final high-stakes question (see Section 3.3). Finally, they were presented with the post-experimental questionnaire eliciting their subjective assessment of the advisors (see Figure 2).

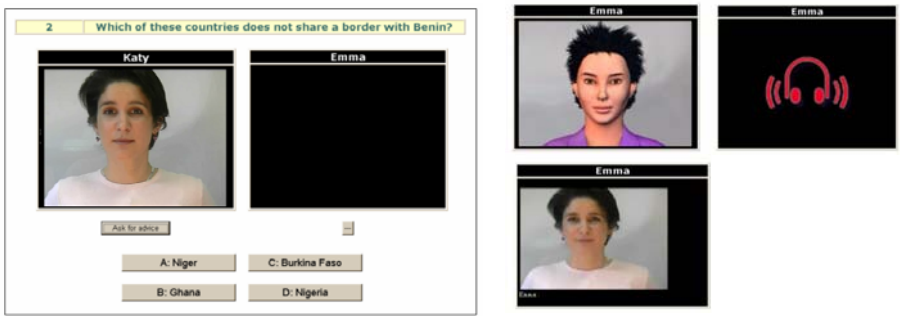


**Figure 1. Experimental system (video advisor selected) and avatar, audio, photo+text.**

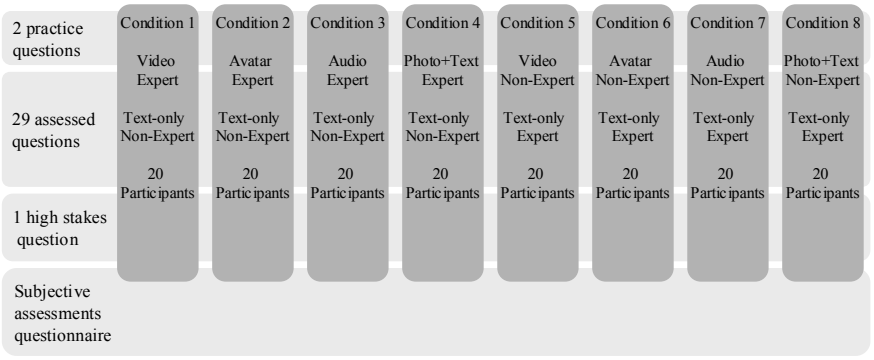


**Figure 2. Overview on the experimental procedure and design.**

## *3.2 Questions*

To minimize effects of participants' prior knowledge, difficult general knowledge questions were used in the quiz study. To choose the 30 most difficult questions out of a pool of 50, an online pre-study was performed with 80 pre-testers who did not take part in the main part of the study. The most difficult question was defined as the one where the two most often picked answers had the smallest difference in their frequency of being chosen. Examples of questions that were included are '*Who coined the term Philosophical Hermeneutics?*' and '*Who won the Turner Prize in 1984?*'. For the 30 questions that were included in the main study, the mean probability for giving a correct answer was .31 (SD = .11), based on the pre-test results. This value is only marginally above chance (.25), indicating that very difficult questions that had been picked.

## 3.3 Independent Variables

**Expertise.** The *non-expert* and *expert* advisors were created by recording advice 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. As each participant only had access to one rich media representation of the advisor, they were unaware that both advisors were in fact the same individual recorded at different levels of *expertise*. In the interest of ecological validity, the phrasing of the advice was not prescribed. Based on experience with a pilot study, 6 incorrect (and less confident) pieces of advice from the untrained recording were added 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.** All media representations were created from the same video clips ranging from 1 sec. to 8 secs. long. 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 behaviour. 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 text appeared dynamically with a delay of 107 ms per letter to ensure that all representations had equal playing time.

**Risk.** Participants' pay was linked to the number of correctly answered questions and thus to their ability to identify the *expert* advisor from interpersonal cues, as the quiz questions were extremely difficult. Pay varied between the £8 and £15. A final high-stakes question (worth an additional £3) was included.

## 3.4 Dependent Variables

### 3.4.1 Advice Seeking

On each question, participants were only allowed to ask one advisor. Seeking advice from one advisor in preference over the other could thus be interpreted as trust in that advisor, as receiving poor advice carried the risk of missing out better advice and therefore reduced participation pay. 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 each participant had two advisors, but could only choose one of them for advice on each question, the following relationships hold: *expert advice seeking = 1 - non-expert advice seeking* and *rich media advice seeking = 1 - text-only advice seeking.*

Figure 3 illustrates P1 and P2 for the measure *advice seeking*. 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.

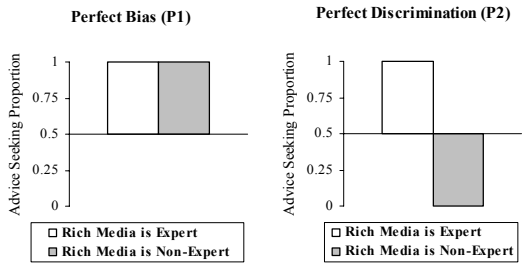


**Figure 3. Illustration of predictions P1 and P2 for advice seeking.**

### *3.4.2 Advice Uptake*

The second measure taken from participants' behaviour was their *advice uptake*, i.e. whether they followed advice they had received from a particular advisor. Again, following advice can be seen as a trusting behaviour, as incorrect answers lead to a lower participation pay. The measure *advice uptake* was defined as the proportion of pieces of advice from one advisor that are followed relative to the total number of times that advisor was asked. While *advice seeking* contrasted P1 and P2 (see Section 3.2.1) within one measure, *advice uptake* gave individual measures for each advisor. Applying the predictions to advice uptake, P1 (media bias), would lead to a higher *advice uptake* for rich media representations, whereas P2 (better discrimination) would lead to a greater effect of expertise on *advice uptake* in richer representations (i.e. an interaction effect between *expertise* and *media representation*).

### *3.4.3 Sensitivity*

To investigate participants' discriminative ability (P2) in different media representations further, their *sensitivity* to correct advice was calculated from the *advice uptake* measure. This measure takes in account the correctness of the advice received (Table 2).

| | **Correct Advice** | **Incorrect Advice** |
|---|---|---|
| **Follow** | Well-placed Trust | Misplaced Trust **(Gullibility Error)** |
| **Not Follow** | No Trust **(Incredulity Error)** | No Trust **(Justified)** |

**Table 2. Correctness of trust decisions (adapted from Fogg 2003).**

Participants had to assess the correctness of a piece of advice from the interpersonal cues they perceived. This can be understood in terms of a sender and receiver model: the advisor's media representation determined the types and number of cues transmitted. Employing a signal detection paradigm (Thurstone 1927), *sensitivity to correct advice* is a measure of the Receiver Operating Characteristics (ROC). The *sensitivity* measure adopted is p(A), a non-parametric

variant of d'(McNicol 1972). p(A) is the area under a ROC-curve (Figure 4), which is defined by the proportion well-placed trust and misplaced trust (Table 2).

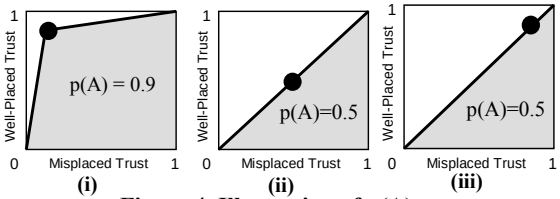


**Figure 4. Illustration of p(A).**

Three examples are illustrated: in Figure 4(i) the participant almost always follows correct advice and almost never follows incorrect advice. In this case the area under the curve, p(A), approaches 1.0 – the participant has a high sensitivity. In Figure 4(ii) the user decides randomly whether to follow advice whether it is correct or not. In this case, the area under the curve p(A) = 0.5, indicating that they cannot detect correct advice (low sensitivity). In the final example, the participant has a tendency to follow any advice given. In this case, p(A) is again = 0.5, as there is no evidence of sensitivity to correct advice. This measure is thus independent from individuals' response biases. Applied to P2, the measure *sensitivity* predicts that richer media result in higher sensitivity scores. P1 predicts no effect on sensitivity.

### *3.4.4 Auxiliary Measures*

As auxiliary measures, participants' self-reports were recorded. For each question, each participant was asked to rate his or her *confidence* in the answer they had given. In addition, participants' subjective assessment of the two advisors was elicited after they had completed the study. Agreement with the statements was elicited on 7-point Likert scales with the anchors 1 ("Strongly disagree") - 7 ("Strongly agree"). In a final open-ended question participants were asked to state the reasons for their advisor choice.

## 4 Results

On average, participants sought advice on 26 out of 30 rounds (87%). Only 51 participants (32 %) sought advice in every round, even though there was no cost associated with seeking advice. One participant (in the *audio expert* advisor condition) did not ask for advice at all. Participants spent on average 23 secs. on each question. If they asked for advice, they did so on average 13 secs. after the question had been displayed, indicating that they first formed their own opinion before asking an advisor.

## *4.1 Advice Seeking*

Figure 5 shows a main effect for *expertise* on participants' likelihood for seeking advice ($F$ (1, 154) = 51.56, $p < .001$). This shows that the experts were chosen

much more often than non-experts for all types of representation. There is also some indication for a between-subjects effect of the type of *rich media* representation ($F$ (3, 154) = 2.50, $p$ = .062). This indicates that the type of *rich media advisor* that was paired with the *text-only* advisor affected how participants decided between the two.

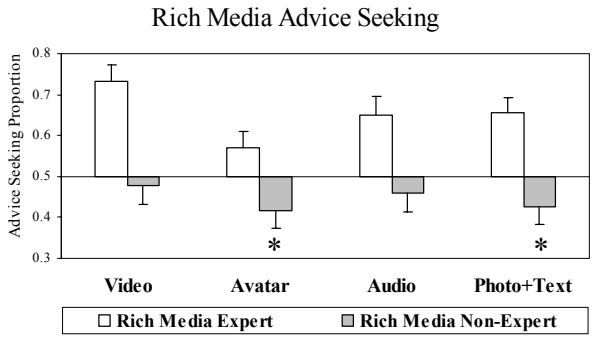


**Figure 5. Seeking advice from the rich media advisor. Stars (*) indicate results for one-sided t-tests (H: *seeking* < .5; $p$ < .05).**

To conduct a within-subject test for bias (P1, Figure 3) and discrimination (P2, Figure 3), we investigated *rich media non-expert* advice seeking (grey bars in Figure 5). As discussed in Section 3.4.1, a value < .5 would provide evidence for discrimination, a value > .5 would be a sign of bias outweighing discrimination. Figure 5 shows *non-expert avatar* and *photo+text* advice seeking significantly below .5 ($t$ (19) = 2.00, $p$ < .05 and $t$ (19) = 1.76, $p$ <. 05, respectively). No such effect is present for *video* and *audio,* indicating that a media bias towards *audio* and *video* is interfering with users' ability to discriminate. In other words, users are *seeking* advice from *video* and *audio* representations equally often, even though they are *non-experts*.

Further evidence for a preference for *seeking video* and *audio* is given by the finding that for *video* and *audio 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; see Figure 6). This effect was not present for the *avatar* and *photo+text* representations. *Avatar expert* advice was sought less often than advice from the other rich media *experts* combined ($t$ (77) = 2.45, $p$ < .05).

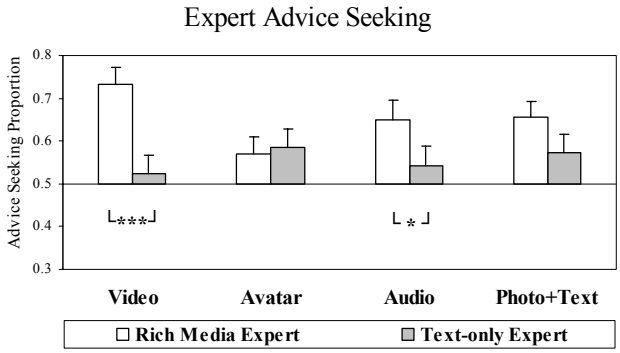


**Figure 6. Advice seeking for the *expert* advisor.**

## *4.2 Advice Uptake*

As stated in Section 3.4.2, a media bias (P1) in *advice uptake* is present if one media representation leads to a higher proportion of uptake than another. Figure 7 shows *advice uptake* for all media and expertise conditions. In addition, it includes aggregate data for the *text-only* advisor, which was present with each of the other media representations (see Section 3.1). In line with the findings for *advice seeking*, the data for *advice uptake* shows a strong effect for *expertise* ($F$ (1,146) = 85.40, $p$ <. 001). In contrast to the findings for *advice seeking*, a between-subject analysis yields no indication of an impact of media representation ($F$ (3,147) = 1.86, *ns*.).

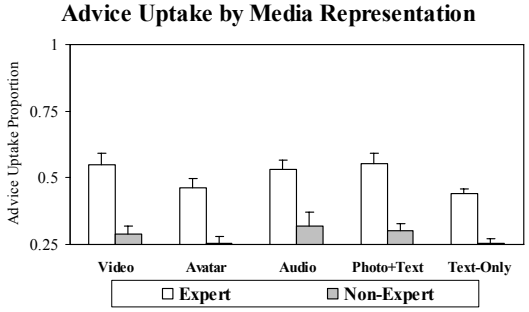


**Figure 7. Advice uptake by media representation and expertise (plotted against .25 which would be random uptake in the presence of 4 answer options, see Section 3.1).**

## *4.3 Sensitivity*

The *sensitivity* measure gives values between 0 and 1. A value =< .5 means that a participant could not differentiate correct from incorrect advice. The *sensitivity* for advice delivered in the different media representations are shown in Figure 8. For each media representation, we tested, whether the sensitivity was greater than 0.5, i.e. whether participants were able to discriminate between correct (confident) and

incorrect (less confident) advice (Table 2). When the *rich media advisor* was an *expert*, participants were sensitive to the differences between correct and incorrect advice. Interestingly, when the *text-only advisor* was the expert (and thus paired with any *rich media non-expert advisor*, see Section 3.1), it also resulted in a *sensitivity* score higher than 0.5. There was no sensitivity for advice given by the *non-expert*.

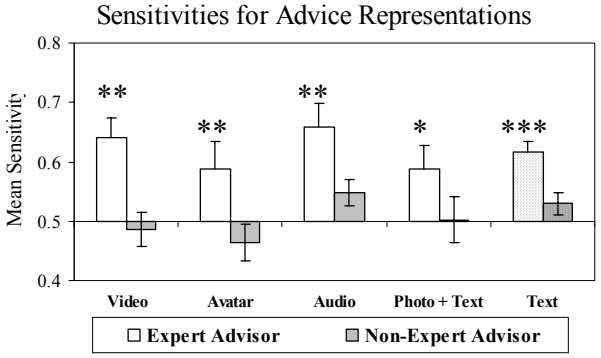


**Figure 8. Sensitivities in experimental conditions. Stars (*) indicate results for one-sided t-tests (H: sensitivity > .5; all text-only advisors collapsed into one bar).**

## *4.4 Auxiliary Measures*

While the focus of this research is on behavioural measures, participants' self-reports were also analysed as auxiliary measures. Getting advice from an *expert* resulted in higher self-reported confidence with an answer ($F$ (1, 154) = 11.76, $p$ < .001), but there was no effect of *media representation* on self-reported confidence. We analysed the post-experimental assessments of the advisors by comparing each participant's rating of the *text-only* advisor to that of the *rich* media advisor, irrespective of the *expertise* of each advisor. Significant differences in assessment between *text-only* and *rich* media advisor are thus indicators of media bias on one statement for one specific *rich* media representation (P1). Notable bias was found for video, which was trusted (S3, Figure 9) more, and rated as being better suited for assessing certainty (S7, Figure 10) than text-only, irrespective of expertise. No such bias was found on these statements for *avatar* and *photo+text* representations. All *rich* media representations resulted in higher ratings of enjoyment (S4; Figure 11).

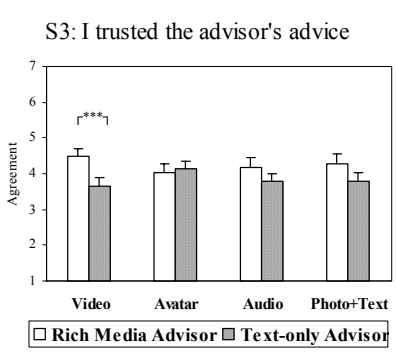


**Figure 9. Self-report for trust in the advisors (S3).**

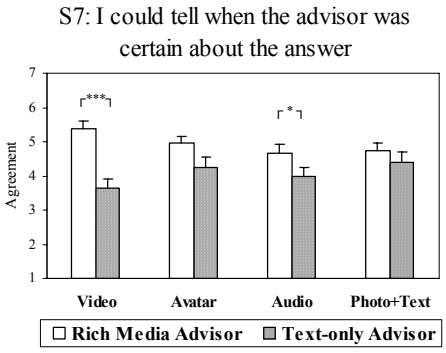


**Figure 10. Self-report for ability to infer advisor certainty (S7).**

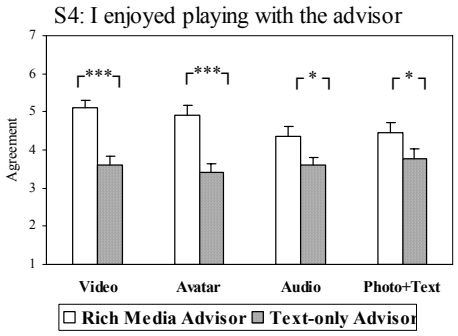


**Figure 11. Self-report of enjoyment of playing with an advisor (S4).**

# 5 Discussion

## *5.1 Video*

When the *non-expert* was represented in *video,* preference for choosing *video* almost matched the preference for choosing *expert text-only* advice. Hence, in many cases users' preference for receiving *video* advice lead them to disregard better *text-only* advice. This preference for video is particularly problematic, as *video* did not result in a higher *sensitivity* for correct advice than *text-only* advice. This finding opposes prediction P2 – that rich media leads to better discrimination. Participants' own post-experimental assessments, however, appear to support P2: they rated their ability to infer certainty (S7) higher for *video* than for *text-only*. One participant expressed this in her reply to the open-ended question:

> *"Since I could see Katie speak and look at her expressions while she answered, I could guess with more confidence when she was correct and thus I chose Katie more number of times."*

This disparity between self-reports and actual performance corroborates a similar finding by Horn et al. (2002) in a study on deception detection over video channels. In that study participants had over-estimated their own ability in detecting lies over video. Horn et al. (2002), in the same study, also found that a severe degradation of the visual channel led to an increase in participants' ability in detecting lies. They hypothesized that the visual identification induces a 'truth bias' that may lead individuals to commit gullibility errors (Table 2). There is no clear indication in the behavioural data of our study for such an effect of the visual channel in particular. However, participants' self-reports suggest a bias resulting from *video* that was not present for other rich media representations. Irrespective of the advisor expertise, participants stated that they trusted the *video* advisor more than the *text* advisor (S3). This effect was only found for *video*, i.e. only in the presence of real dynamic visual interpersonal cues.

In summary, participants were able to identify expert advice in the *video* representations, but the data suggest that the additional cues received in *video* compared to *text-only* did not increase their *sensitivity* to correct advice (P2); rather, there is some evidence (in *advice seeking* and in the self-reports) that participants had a tendency to trust *video* (P1), which interfered with their ability to detect expertise.

## *5.2 Audio*

Similar to *video*, the preference for seeking *non-expert audio* advice almost matched the preference for *expert* advice, which indicates that the tendency for seeking *audio* advice interfered with participants' preference for *expert* advice (P1). Participants over-estimated their ability to detect certainty (S7) in audio, as they did for video. However, unlike for *video*, participants did not state that they trusted the *audio* advisor more than the *text* advisor (S3). *Expert audio* advice resulted in a good sensitivity to correct advice, but it was not significantly better than the sensitivity in the *text-only* or any other media representation (P2). In summary, there is some behavioural evidence for interference from a preference for audio representations on users' ability to discriminate, but on users' self-reports less bias was detected for audio than for video.

## *5.3 Avatar*

The *avatar* did not result in a positive bias. To the contrary, avatar expert advice was less often sought than other types of *rich media expert* advice. The subjective assessments corroborate the notion of a negative bias resulting from the avatar: participants did not think it had been easier to assess the *avatar's* certainty relative to the *text-only* advisor's (S7). For the *audio* advisor, which provided the same audio cues, but not the synthetic visual ones, they considered themselves able to do so. In the words of one participant:

> *"Katy didn't seem real so I stopped picking her for advice."*

Clearly, these findings cannot necessarily be generalized to other avatar representations or contexts of use. Previous studies showed strong differences in reactions to animated characters due to relatively small differences in behaviour, appearance, or context of use. Nonetheless, the findings indicate that using an avatar created with off-the-shelf tools from the audio stream may not be advisable for building trust at this stage. If the avatar and all the visual interpersonal cues given off had been carefully scripted, the *avatar* advisor might have appeared to be more trustworthy. Finally, it was found that the *avatar* was perceived as more enjoyable (S4) than the *text-only* advisor. Our data thus suggest that this representation may be effectively be used in e.g. an entertainment context.

## 5.4 Photo + Text

The *photo+text* advisor offered the fewest additional interpersonal cues relative to the *text-only* advisor. The static visual cues given in a photo did not carry any information about expertise or confidence of individual pieces of advice. Hence, this representation could not be expected to increase participants' ability to discriminate between advisors based on their expertise (P2). Only a bias (P1) arising from the presence of static interpersonal cues, could be expected. No such bias was found on any of the measures, but the *photo* did result in higher ratings for enjoyment (S4) compared to *text-only*. This suggests that photos can be used to prompt positive reactions and make interactions more engaging.

## 5.5 Media Bias (P1) and Discrimination (P2)

Averaging across all rich media representations there was strong evidence for a preference for seeking expert advice and some evidence of media bias (P1). Users' ability to discriminate between expert and non-expert was good, independent of the media representation (P2).

Investigating the rich media representations individually, it was found that the preference for seeking *expert* advice was almost matched by a preference for seeking advice in *video* and *audio* representations. However, this effect was not so strong as to supersede preference for *expert* advice. Nonetheless, it led to participants receiving less trustworthy advice than they otherwise would have. In other words, their preference for video and audio led them to disregard good advice that was given as *text-only*.

Participants' self-reports show that they preferred to seek *video* and *audio* advice, because they thought these representations allowed them to make better trust assessments of individual pieces of advice: they considered their ability to infer advisor certainty in *video* and *audio* representations as higher than in text-only (S7). However, there is also evidence for media bias (P1) in participants' self-reports: for *audio* and *video* participants stated that they trusted the rich media advisor more than the *text-only* advisor (S3).

Whereas effects on *advice seeking* behaviour and users' self-reports are important, one could argue that the real test for media bias is whether someone acts

on advice. Hence, we also investigated *advice uptake*. While there was an effect of advisor expertise on advice uptake, no effect of media representation, i.e. no media bias was found (P1). This result is reassuring as it shows that users' trust, measured by *advice uptake*, cannot be easily swayed by choice of media representation. On the other hand, the lack of effect of media representation also showed that the richer representations *video* and *audio*, which participants evidently – based on their *advice seeking* behaviour and self-reports – considered as giving more insight into trustworthiness, did not allow an improved discrimination between trustworthy and less trustworthy pieces of advice. This conclusion is also supported by the results of the *sensitivity* measure, which did not investigate participants' ability to detect the expert, but their sensitivity to the correctness of individual pieces of advice. There was no significant difference in *sensitivity* to correct advice between all the *rich* media representations and *text-only*. The good performance at detecting the correctness of expert advice in all media representation suggests that most information was conveyed in lexical cues and that little extra information could be gained by the other interpersonal cues (para-verbal and visual cues) that were conveyed in the rich media representations. Hence, while participants thought that *audio* and *video* offer them superior *sensitivity*, no such effect was found.

## 6 Conclusions

This study investigated whether richer representations result in either a positive media bias (P1) or increased sensitivity for cues of expertise (P2) compared to *text-only* representations. We analysed participants' *advice seeking* and their *sensitivity* to correct advice in a situation of limited advice and financial risk. We found that participants mainly *sought advice* from the expert advisor, irrespective of the media representation (P2). However, we found no *sensitivity* for correct advice when the advisor was a *non-expert*. This indicates that participants could not identify subtle differences between low levels of confidence in any media representation. For *expert* advice, participants showed higher *sensitivity* in all representations, including *text-only*. This finding questions classic media richness models that predicted that text-only communication suppresses cues that are essential for trust assessments.

Results for participants' advice seeking suggest that a bias (P1) for *audio* and in particular *video* representations can interfere with users' ability to discriminate effectively. The interference was caused by users' belief in the superiority of these media for trust assessments, which mirrored classic media richness models. This belief led them to choose *audio* or *video* over *text-only* even at the cost of missing out on *expert* advice. This preference could have negative consequences for users. Consider, for instance, a user browsing a health advice site and focusing exclusively on video advice – and thereby missing our on potentially better text advice. Hence, for designers, who wish to ensure high levels of trust, video is the best representation, followed by audio.

The *avatar* was not found to have a positive effect on trust. Using an avatar without careful scripting from an audio stream may not be an advisable strategy for building trust with the current state of art of avatar development. However, 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 inducing trust, our data suggests that these representations can be effective.

In this study we introduced a measure from signal detection theory, p(A), to assess participants' ability to place trust correctly. As it is an easily calculated measure that captures both, (1) correctly placed trust and (2) correctly withheld trust, it can be employed in future studies, which seek to manipulate trustworthiness to assess the correctness of trust decisions. Since we found disparities between participants' self-reports and their actual behaviour, our results also provide further support for measuring trust by observing decision-making under risk, rather than only relying on self-reports.

Whilst this study exclusively looked at cues for expertise in the context of a general knowledge quiz, future studies could usefully employ a similar paradigm to research media effects for cues of motivation (e.g. wilful deception) in different trust-requiring situations.

## Acknowledgements

We would like to thank Cyril Scott at DA Group (www.dagroupplc.com) for the V1 animation tool, as well as M. Garau, P. Bonhard, H. Knoche, Alan Johnston, Clare Harries (UCL), S. Mahlke (TU Berlin), Richard Boardman (Google), and K. Chorianopoulos (IC London). We would further like to thank the anonymous reviewers for their helpful comments.

## References

Bacharach, M. and Gambetta, D. (2003). "Trust in Signs". In: Cook, K. S. *Trust in Society*. Russell Sage: New York, NY.

Bos, N., Olson, J. S., Olson, G. M., Wright, Z., and Gergle, D. (2002). "Rich media helps trust development". *Proceedings of CHI2002*. ACM Press: New York, 135-140.

Corritore, C. L., Kracher, B., and Wiedenbeck, S. (2003). "On-line trust: concepts, evolving themes, a model". *International Journal of Human Computer Studies* **58**(6), 737-758.

Davis, J. P., Farnham, S. D., and Jensen, C. (2002). "Decreasing online 'bad' behavior". *Extended Abstracts CHI 2002*. ACM Press: New York, 718-719.

Deutsch, M. 1958. "Trust and suspicion". *Journal of Conflict Resolution* **2**(3), 265-279.

Fogg, B. J. (2003). *Persuasive Technology*. Morgan Kaufmann: San Francisco.

Giddens, A. (1990). *The consequences of modernity*. Stanford University Press: Palo Alto.

Goffman, E. (1959). *The Presentation of Self in Everyday Life*. Doubleday: Garden City.

Handy, C. (1995). "Trust and the Virtual Organization". *Harvard Business Review* **73**(3), 40-50.

Hinton, P. R. (1993). *The Psychology of Interpersonal Perception.* Routledge: London.

Horn, D. B., Olson, J. S., and Karasik, L. (2002). "The Effects of Spatial and Temporal Video Distortion on Lie Detection Performance". *CHI 2002 Extended Abstracts.* ACM Press: New York, NY, 716-718.

McNicol, D. A. (1972). *A Primer of Signal Detection Theory.* Allen & Unwin: London.

Olson, J. S., Zheng, J., Bos, N., Olson, G. M., and Veinott, E. (2002). "Trust without Touch: Jumpstarting long-distance trust with initial social activities". *Proceedings of CHI2002*. ACM Press: New York, NY, 141-146.

Rice, R. E. (1992). "Task analyzability, use of new medium and effectiveness: A multi-site exploration of media richness". *Organization Science* **3**(4), 475-500.

Short, J., Williams, E., and Christie, B. (1976). *The Social Psychology of Telecommunications*. John Wiley & Sons: London.

Swerts, M., Krahmer, E., Barkhuysen, P., and van de Laar, L. (2004). "Audiovisual cues to uncertainty". *Proceedings of the ISCA Workshop on Error Handling in Spoken Dialogue Systems*, Chateau-D'Oex, Switzerland, 2003.

Thurstone, L. L. (1927). "Law of Comparative Judgement". *Psychological Review* **34**, 273-286.

van Mulken, S., Andre, E., and Müller, J. (1999). "An empirical study on the trustworthiness of life-like interface agents". *Proceedings of HCI International '99, 2*, Lawrence Erlbaum: Mahwah, 152-156.

Walther, J. B. (1999). "Visual Cues and Computer Mediated Communication: Don't look before you leap" *Proceedings of the Annual Meeting of the ICA,* San Francisco, CA, US.

Whittaker, S. and O'Conaill, B. (1997). "The Role of Vision in Face-to-Face and Mediated Communication". In: Finn, K. E., Sellen, A. J., and Wilbur, S. B. *Video-Mediated Communication*. Lawrence Erlbaum: Mahwah, NJ, 23-49.

Winston, J. S., Strange, B., O'Doherty, J., and Dolan, R. J. (2002). "Automatic and intentional brain responses during evaluation of trustworthiness of faces." *Nature Neuroscience* **5**(3), 277-283.

Yaniv, I. and Kleinberger, E. (2000). "Advice taking in decision making: Egocentric discounting and reputation formation." *Organizational Behavior and Human Decision Processes* **83**, 260-281.

Zuckerman, M., DePaulo, B. M., and Rosenthal, R. (1981). "Verbal and nonverbal communication of deception". *Advances in Experimental Social Psychology* **14,** 1-59.