

# Using Digital Images to Enhance the Credibility of Information

Hien Nguyen<sup>1</sup> and Judith Masthoff<sup>1</sup>

**Abstract.** With research revealing the importance of trust as a deciding factor for users to visit a website, designing websites that users find credible becomes important. Adding onscreen characters in various forms (e.g. photographs of a person) to represent the source of information is a simple and popular way to increase the credibility of a website. However, despite its popularity, empirical studies have reported mixed results. This paper presents three experiments that explored the causes of this contradiction. In our experiments, the credibility of the source (as resulting from a photograph of a person) was found to be topic dependent. It was also found to positively correlate with and account for some 30% of the variability in the credibility of a website. These findings perhaps can explain the inconsistency of many previous studies. Finally, our study suggested that adding a user-selected photograph increases the credibility of the website, while the opposite occurs when adding a photograph that is lowly credible with respect to the topic of the website's content.

## 1 INTRODUCTION

Social psychology has long emphasized the importance of the credibility (i.e. being trustworthy and knowledgeable) of the source (of information) in a persuasive communication. The credibility of the source is not a commodity that the source possesses, but it is the perception of the receiver about the source [1]. The more credible the receiver finds the source, the more credible the messages conveyed by the source become and vice versa [1]. It is acknowledged that one way used by the receiver to judge the source's credibility is through the source's appearance [2]. The source's appearance influences its credibility by affecting the receiver's perception of the underlying factors that form credibility. These factors commonly include *physical attractiveness*, *trustworthiness*, and *expertise* [1,2].

One way that web designers have tried to utilize this suggestion to enhance the credibility of (the information of) websites is by using onscreen characters to deliver the sites' information. This could be as simple as having a photograph of a person attached to the information, or as complex as having a fully animated character that is capable of exhibiting life-like behaviours such as speech, gestures, or expressions of emotions. However, despite its popularity, empirical studies have reported mixed results of this approach. While some studies highlighted the intended positive effect, others suggested otherwise. Two studies at Boston University showed that people were more

willing to cooperate with a human-like character when that character had been made more attractive [3]. However, attractiveness alone was not sufficient to predict cooperation: subjects cooperated less with a more attractive, but dog-like character. Another study showed that adding a formal photograph of an author improved the trustworthiness, believability, perceived expertise, and competence of a web article (compared to an informal or no photograph) [4]. However, adding a photograph of a person did not enhance the perceived trustworthiness of a computer's recommendations [5]. Similarly, while Steinbrueck et al. [6] found that images of employees can increase users' trust in e-commerce sites, a later study by Riegelsberger [7] could not replicate this result. In his study, "neither the presence of a photo, nor the trustworthiness of the person depicted, had a significant effect" [7]. Other studies reported both positive and negative effects (e.g. [8]), or positive effects in some situations but not others (e.g. [9]). So what causes this inconsistency?

In this paper, we thoroughly explore this contradiction, focussing on one particular form of onscreen characters: *photographs of people*. Section 2 presents our hypotheses, while Section 3 discusses our experiments and their results with respect to the hypotheses postulated in Section 2. Finally, Section 4 summarises our findings and their implications.

## 2 RESEARCH QUESTIONS

In our earlier attempt to solve this contradiction [10], we argued that one reason for the inconsistency discussed in Section 1 is that the credibility of a person may well be topic dependent. When delivering information, a speaker might have high credibility in certain topics and low credibility in others. For instance, we can look at a user who is trying to be more physically active. In the user's situation, s/he might find a doctor more credible than a sport instructor to talk to about the benefits of exercise on health. In contrast, a sport instructor might have an edge over a doctor when talking about fitness programmes. Meanwhile, a close friend or someone who is similar to the user and has been in the user's situation might be the most credible person should the user need social support. This prediction led to our first hypothesis:

**H1:** *The (perceived) credibility of a person (as resulting from his/her appearance) is topic dependent.*

Indeed, H1 was confirmed by Experiment 1 in [10]. So perhaps, we did not see the expected results in many aforementioned studies because the selected photos were not highly credible with respect to the particular topic of the website that they were

---

<sup>1</sup> Dept. of Computing Science, Univ. of Aberdeen, AB24 3UE, UK.  
Email: {n.nguyen, j.masthoff}@abdn.ac.uk;

attached to (e.g. [5,7]), or the same photo was used across many websites promoting different products (e.g. [7,9]).

Therefore, we would expect that adding a photograph that is highly credible with respect to the topic of a website improves the credibility of the website. In our earlier study (see Experiment 2 in [10] for a complete discussion), we asked participants to judge the credibility of two web pages, one after the other. The first page informs the users about the health benefits of exercise, the second page provides guidance on planning personal fitness programmes. The web pages were presented to the participants with either a highly credible photo (HI, 19 participants), a lowly credible photo (LI, 19 participants) or without a photo (21 participants). Similar to other studies mentioned in Section 1, the credibility of the photographs was validated by a number of subjects prior to the experiment, and these subjects did not take part in the experiment.

	Source Credibility	Content Credibility
A web page about the health benefits of exercise		
HI (photo 7)	3.95	4.62 (0.84)
LI (photo 4)	3.02	3.90 (0.84)
No photo	-	4.30 (0.90)
A web page about planning personal fitness programmes		
HI (photo 12)	3.87	4.37 (0.74)
LI (photo 8)	3.25	4.07 (0.73)
No photo	-	4.32 (0.76)

**Table 1.** The results of our earlier study. Reported in [10].

The results of the experiment are summarised in Table 1. We found that attaching a photo with a high credibility increases the credibility of the first page that discusses the health benefits of exercise, while attaching a photo with a low credibility decreases its credibility (compared to the credibility of the page when no photo was present). The difference between the groups was statistically significant. The same trend was found for the second page that discusses how to plan personal fitness programmes. However, the magnitude of the increase and decrease of the credibility of the page were not statistically significant in this case. Some explanations could be offered to justify this result. First, the gap between the credibility of the two photos attached to the second page (0.62) was smaller than that of the two photos used in the first page (0.93), and this might have weakened the effect they had on the credibility of the page. Secondly, participants may not have taken the second part of the experiment as seriously, due to tiredness. Finally, it might well be that the participants who took part in the experiment did not judge the credibility of our highly credible photo as high, and our lowly credible counterpart as low as we had validated. Since credibility is a perception, it is unlikely that all viewers will perceive a particular photograph of a person equally. However, the right trend in our results suggests a positive correlation between the credibility of the photograph and the credibility of the content of the web page. This leads to our second hypothesis:

**H2:** *The perceived credibility of the person shown in the photograph with respect to a topic is positively related to the perceived credibility of the information of the website on that topic.*

If the credibility of a person is topic and receiver dependent as we pointed out, and if the credibility of a person on a topic positively correlates with the credibility of the information on that topic, then a user should find a message more credible if it is conveyed by a person that s/he personally finds highly credible with respect to that topic and vice versa. We also argue that, the user, when given the choice, will select the person whom s/he personally finds the most credible. This leads to our third hypothesis:

**H3:** *The perceived credibility of a message accompanied by a photograph chosen by the user for that topic is higher than that of the same message with no photograph, which in turn is higher than that of the same message with a lowly credible photograph.*

### 3 RESEARCH METHODOLOGY

#### Research Design

Our hypotheses are verified through a series of between-subjects experiments. Participants took part in the experiments voluntarily and no participant took part in more than one experiment.

#### Experiment 1: Source Credibility and Topic

##### Experimental Design

The aim of this experiment is two-fold: (1) to see whether the credibility of a person as judged from a photograph is indeed topic dependent, and (2) to validate the perceived age, gender, profession, and credibility for various potential photographs. While the former aims to test hypothesis 1, the later aims to find a subset of photographs with good inter-subject agreement on all these criteria to be used in future experiments. Details of the experiment were reported in [10], we will only briefly discuss this experiment here to enhance the continuity and understandability of this paper.

	5	4	3	2	1
<b>Attractiveness</b>					
Attractive					Unattractive
Classy					Not classy
Beautiful/Handsome					Ugly
Elegant					Plain
Sexy					Not sexy
<b>Trustworthiness</b>					
Dependable					Undependable
Honest					Dishonest
Reliable					Unreliable
Sincere					Insincere
Trustworthy					Untrustworthy
<b>Expertise</b>					
Expert					Not an expert
Experienced					Inexperienced
Knowledgeable					Unknowledgeable
Qualified					Unqualified
Skilled					Unskilled

**Table 2.** Ohanian's credibility scale of a person.

In the experiment, the *credibility* of a person was defined as the average of his/her *attractiveness*, *trustworthiness*, and *expertness*, since these three factors are widely acknowledged to have the strongest influence on credibility [1,2]. A person's attractiveness, trustworthiness, and expertness were measured using 15 5-point Semantic Differential scale items developed by Ohanian [11] to measure a person's credibility. Each factor is calculated as the average of its five corresponding items (see Table 2 for exact wordings).

All participants were presented with 16 head and shoulder photographs of doctors and sport instructors / athletes. All photographs were taken from Microsoft Clipart (using search keywords like doctor and sport) and varied in age, and gender. The presentation order of the photographs was randomized for each participant to control for order effects. Examples of the photographs are shown in Figure 1. All the photographs can be found in [10].



**Figure 1.** Examples of photographs used in the experiments.

Participants were randomly assigned to one of four groups (to limit the time needed to perform the experiment and to avoid interaction effects between the questions). We asked each group to judge one or more characteristics of the person who appeared in each photograph, namely:

- Group A: *gender* (male or female), *age* (< 25, 25-30, 30-40, 40-45, or > 45), and *most likely profession* (doctor, sport instructor, or other),
- Group B: *attractiveness*
- Group C: *trustworthiness*
- Group D: *expertness* with respect to (1) the health benefits of exercise, and (2) fitness programmes.

Finally, all participants were presented with a web page showing all 16 photographs (the order of the photographs was also randomized for each participant). They could hover over each thumbnail to see the full size version. They were asked to choose whom they would like the most to learn from about each topic (i.e. health benefits of exercise and fitness programmes).

### Participants

Fifty-one participants took part in the experiment (see Table 3 for the distribution of the participants' age and gender). Participants were staff and graduate students of the university, but came from all areas and professions.

	Gender		Age			
	F	M	18-20	21-24	25-29	30+
Number of subjects	32	19	5	17	9	20

**Table 3.** The distribution of participants' age and gender.

### Results

Table 4 shows the average attractiveness, trustworthiness, and expertness of each photo. Hypothesis 1 stated that a person's credibility (as resulting from his/her appearance) is topic dependent. Indeed, all doctors were rated to have more expertise in the health benefits of exercise than in fitness programmes, while the opposite holds for all sport instructors, and two people who were identified as neither doctor nor sport instructor were rated to have the same level of expertise with respect to the two topics (also noted that their expertise was lower than that of both doctors and sport instructors). However, the difference between the levels of expertise in the two topics for sport instructors was much less significant than that of doctors. Perhaps, sport instructors are assumed to be interested in exercise not just for the sake of exercise, but also because they care about its benefits to health. Independent sample t-tests indicated that doctors are perceived as more expert with respect to the health benefits of exercise than sport instructors (average=3.92, stdev=0.21 vs. average=3.59, stdev=0.18,  $p < 0.05$ ), while sport instructors are perceived as more expert with respect to fitness programmes than doctors (average=3.86, stdev=0.22 vs. average=3.26, stdev=0.34,  $p < 0.05$ ). Thus, we concluded that H1 is supported in [10].

Photo	A	T	H Exp	#H	E Exp	#E
Photos which are perceived as doctors						
1	2.9 (0.7)	3.9 (0.8)	4.1 (0.6)	1	3.0 (0.9)	1
2	3.0 (0.9)	4.1 (0.7)	3.8 (1.1)	7	3.0 (0.8)	1
3	3.4 (0.6)	4.1 (0.8)	4.1 (0.6)	1	3.5 (0.5)	1
4	2.4 (0.6)	3.2 (0.9)	3.5 (1.3)	1	2.5 (1.1)	0
5	2.7 (0.6)	4.6 (0.5)	4.0 (0.8)	1	3.4 (0.7)	1
6	2.7 (0.8)	3.7 (0.7)	3.9 (0.7)	2	3.5 (0.6)	3
7	3.3 (0.7)	4.5 (0.5)	4.1 (0.8)	8	3.5 (0.6)	2
8	2.5 (0.7)	3.8 (1.0)	4.0 (0.9)	1	3.4 (0.6)	0
11	3.0 (0.8)	4.1 (0.7)	3.9 (0.6)	1	3.5 (0.7)	0
Photos which are perceived as sport instructors / athletes						
9	3.8 (0.8)	3.2 (1.1)	3.7 (0.6)	5	3.9 (0.5)	14
10	2.8 (0.9)	3.0 (0.9)	3.6 (1.0)	6	4.0 (0.7)	9
12	3.6 (0.6)	3.9 (0.9)	3.8 (0.7)	5	4.1 (0.5)	12
13	2.8 (0.9)	4.3 (0.5)	3.3 (0.9)	1	3.7 (0.9)	2
15	2.6 (0.9)	4.3(0.6)	3.5 (0.8)	7	3.6 (0.8)	3
Photos which are perceived as other						
14	2.4 (0.5)	3.3 (1.2)	3.0 (0.9)	3	3.1 (1.1)	2
16	2.5 (0.9)	3.8 (0.9)	3.2 (1.2)	1	3.2 (1.2)	0

\* A = Attractiveness; T = Trustworthiness.

H Exp, E Exp = Expertness on the topic of health benefits of exercise, fitness programmes respectively.

#H, #E: number of subjects who picked this photo as their favourite to learn from about the health benefits of exercise and fitness programmes respectively.

**Table 4.** Average score and standard deviation of each photo.

## Experiment 2

### Experimental Design

As we previously discussed in section 2, there ought to be a positive correlation between the credibility of a web page and the credibility (with respect to the topic of the web page) of the photo attached to it (H2). This experiment aims to verify whether this correlation does in fact exist.

This experiment is an improvement of our earlier study that we discussed in section 2 (also see [10, Section 3] for a complete discussion). All participants were shown the same web page on how to plan a personal fitness programme that was used in the previous study. We composed the content of the web page based on the information available on [www.mayoclinic.com](http://www.mayoclinic.com)<sup>2</sup>. Participants were asked to read the given web page carefully. After reading the page, they were asked to judge its credibility. The credibility of the page was measured using 15 7-point Likert scale items. The items were developed and validated by Hong [12] for assessing the credibility of health-related websites. They assess five commonly recognised dimensions of the *credibility* of information: *expertise*, *goodwill*, *trustworthiness*, *depth*, and *fairness* (see Table 5 for exact wordings). Credibility is calculated as the average of the five dimensions. Each dimension is the average of its respective items. The items were ordered such that no two items from the same dimension appeared sequentially.

Fairness
This page provides information that is neutral
This page provides information that is <b>not</b> balanced
This page is biased in the information it provides
This page is slanted in the information it provides
This page is even-handed in presenting information
Depth
This page does <b>not</b> provide in-depth information
This page is <b>not</b> comprehensive
This page offers everything you need to know on the topic
Goodwill
This page has my interests at heart
This page is uncaring about its visitors
This page is <b>not</b> concerned about its visitors
Trust/Expertise
This page appears to have experts on the topic discussed
This page is ethical
This page appears to be a leader in its area of specialty
This page is <b>not</b> trustworthy



**Table 5.** Hong's credibility scale for a health-related web page.

All participants also indicated the extent to which they already knew the information presented to them, their knowledge on the topic, and their reliance on the Internet for seeking health-related information. The exact wordings and results are shown in Table 7. Note that the first three questions were asked before, and the last two after the participants read the message.

Participants were randomly assigned to one of three groups. The message for the first (the Highly Credible Image (HI)) and second (the Lowly Credible Image (LI)) group prominently

<sup>2</sup> The content of the web page can be found at [www.csd.abdn.ac.uk/~hnguyen/aisb08/fitness.txt](http://www.csd.abdn.ac.uk/~hnguyen/aisb08/fitness.txt)

showed a highly credible photo (photo 12) and a lowly credible photo (photo 8) respectively. The photos were the same as the ones used in our earlier study (see Table 6). The participants were told that the source of the content of the page was the person who appeared in the photograph, and asked to judge the credibility of the person after they judged the credibility of the page. The credibility of the person was measured in the same way used in Experiment 1. The message for the third group (the No Image (NI)) did not have any photo attached. Participants in this group were also asked to judge the credibility of the source of the message after they judged the credibility of the message. For the NI group, we used the same measures used in Experiment 1 except that the questions about attractiveness were taken out since they were not applicable.

		
	Photo 12	Photo 8
Attractiveness	3.63	2.52
Trustworthiness	3.85	3.83
Expertness in fitness programs	4.13	3.38
<b>Overall Credibility for fitness programmes</b>	<b>3.87</b>	<b>3.25</b>

**Table 6.** The photographs used in the experiment.

### Participants

Thirty-one participants took part in the experiment (see Table 7 for the participants' demographics).

	HI		LI		NI	
Number of participants	11		11		9	
Gender	M	F	M	F	M	F
	5	6	7	4	5	4
Average age	30.00 (15.88)		34.27 (8.43)		30.11 (11.86)	
I am currently doing some form of exercise	Y	N	Y	N	Y	N
	6	5	8	3	8	1
I am more educated about my health than most people*	4.91 (0.83)		4.91 (1.30)		4.44 (1.74)	
I have full knowledge of the benefits (consequences) of regular exercise (the lack of it)*.	5.55 (1.04)		5.45 (1.44)		5.56 (1.88)	
When I need information about benefits of exercise and fitness programmes, I would go to the Internet*.	4.73 (1.68)		4.91 (1.22)		4.33 (2.06)	
I already knew all the information presented**	5.18 (1.08)		4.91 (1.51)		4.89 (2.32)	

\* 1 = strongly disagree 7 = strongly agree

\*\* 1 = nothing 7 = everything

**Table 7.** The participants' demographics.

Participants were recruited online from an experiment listing website maintained by the University of Zurich

(<http://genpsylab-wexlist.unizh.ch/>), and from a mailing list that staff and students of the university subscribed to. Participants came from all areas and professions.



Hello Hien Nguyen.

As you know, fitness programs can be a great way to profit from health benefits of regular exercise. Combining camaraderie and upbeat music with strength training, aerobic dancing or yoga can make for an enjoyable productive workout. But between step aerobics, Pilates, and other trendy programs, **how do you know which fitness program is best for you?**

Perhaps you already know that there's no one fitness program that's perfect for everyone. But with so many options available — from aquatic aerobics to strength training to stationary bicycling — you're certain to find a well rounded routine. These simple steps will help you put your program together.

**Step 1. Determine your fitness goals**

Start by thinking about why you want to start a fitness program. For instance, your doctor may have suggested that you start an exercise program to lose weight. Or, if you're already active, you may want to prepare for a 5K race or get ready for the downhill skiing season. A fitness program that compliments your goals will help you stay motivated.

**Step 2. Identify your strengths and weaknesses**

For example, if you work out on an elliptical trainer at home, but you aren't very flexible, a class that emphasizes flexibility, such as yoga, may be an option. Do you lift weights at home? Then consider signing up for an aerobics class to balance your fitness program. Another thing to consider is your health. Do you have arthritis? Then think about aquatic exercise, which minimizes stress on your joints.

**Step 3. Assess your fitness level**

One option is to meet with a fitness professional. After your assessment, he or she will work with you to pinpoint areas that need attention. You'll find out what your baseline fitness level

**Seven steps to your personalized fitness program**

1. Determine your fitness goals
2. Identify your strengths and weaknesses
3. Assess your fitness level
4. Consider your likes and dislikes
5. Weigh your options
6. Get started
7. Monitor your progress

**Seven steps to your personalized fitness program**

1. Determine your fitness goals
2. Identify your strengths and weaknesses
3. Assess your fitness level
4. Consider your likes and dislikes
5. Weigh your options
6. Get started
7. Monitor your progress

Hello Hien Nguyen.

As you know, fitness programs can be a great way to profit from health benefits of regular exercise. Combining camaraderie and upbeat music with strength training, aerobic dancing or yoga can make for an enjoyable productive workout. But between step aerobics, Pilates, and other trendy programs, **how do you know which fitness program is best for you?**

Perhaps you already know that there's no one fitness program that's perfect for everyone. But with so many options available — from aquatic aerobics to strength training to stationary bicycling — you're certain to find a well rounded routine. These simple steps will help you put your program together.

**Step 1. Determine your fitness goals**

Start by thinking about why you want to start a fitness program. For instance, your doctor may have suggested that you start an exercise program to lose weight. Or, if you're already active, you may want to prepare for a 5K race or get ready for the downhill skiing season. A fitness program that compliments your goals will help you stay motivated.

**Step 2. Identify your strengths and weaknesses**

For example, if you work out on an elliptical trainer at home, but you aren't very flexible, a class that emphasizes flexibility, such as yoga, may be an option. Do you lift weights at home? Then consider signing up for an aerobics class to balance your fitness program. Another thing to consider is your health. Do you have arthritis? Then think about aquatic exercise, which minimizes stress on your joints.

**Step 3. Assess your fitness level**

One option is to meet with a fitness professional. After your assessment, he or she will work

**Figure 2.** The website with and without photo.

**Results**

Hypothesis 2 stated that the perceived credibility of the person shown in the photograph with respect to a topic positively correlates with the perceived credibility of the information of the website on that topic. Table 8 shows the Pearson correlation test results.

Overall, the credibility of the source is positively related to the credibility of the message when a photograph of the source is present ( $r=0.557$ ,  $R^2=0.310$ ,  $p<0.01$ ). This result also pointed out that the credibility of the source accounts for some 31% of the variability in the credibility of the message, hence it would not be the only factor influencing the users' perception of the message credibility, especially when considering how simple our site was (see Figure 2).

This might give a good explanation for the non-significant correlation between the source and the message credibility found in the NI and LI group. We noted that the majority of the participants in LI and NI groups are currently exercising, which was not the case for the HI group. This may have weakened the impact of the credibility of the source on the credibility of the message, leading to the non-significant correlation. A follow-up experiment will be needed to investigate these issues further.

		Source			Overall credibility
		Attractiveness	Trust	Expertise	
Web page credibility	HI	0.143	0.853*	0.733*	<b>0.695*</b>
	LI	0.098	0.395	0.296	<b>0.313</b>
	With photo (HI+LI)	0.125	0.710*	0.544*	<b>0.557*</b>
	NI	n/a	0.554	0.233	<b>0.446</b>

\* Correlation is significant at the 0.01 level (1-tailed)

**Table 8.** Pearson's correlations between the credibility of the source and the credibility of the web page.

**Experiment 3**

**Experimental Design**

This experiment is designed to verify our third hypothesis, which stated that adding a user-selected photograph increases, while adding a lowly credible photograph decreases the perceived credibility of the information.

All participants were asked to select the reasons why they find regular exercise difficult from a set of ten common reasons (e.g. "I don't have enough time to exercise", "I'm too tired to exercise after working all day"). For each reason, the participants can rate its importance on a 5-point Likert scale (from "not at all" to "strongly agree"). Participants were then presented with either a standard or personalised message containing suggestions to their problems (e.g. "Claim the back row of the parking lot as your own. Or park a few blocks away and walk quickly to your destination"). They were asked to read the given information carefully. They were then asked to judge the credibility of the message and the source. These were set up in the same way used in Experiment 2.

All participants also indicated the extent to which they already knew the information presented to them, their knowledge on the topic, and their reliance on the Internet for seeking health-related information. The questions were the same as those used in Experiment 2.

Participants were randomly assigned to one of four groups:

- (1) Personalised Image and Personalised Content (PI+PC).
- (2) Lowly Credible Image and Personalised Content (LI+PC).
- (3) No Image and Personalised Content (NI+PC).
- (4) No Image and No Personalised Content (NI+NPC).

Participants in the first group (PI+PC) were asked to choose someone whom they would like to speak to by clicking on one of the sixteen photos used in Experiment 1, prior to selecting reasons for finding regular exercise difficult. They were then presented with a personalised message. The chosen image was prominently shown alongside the message. The message was personalised to only contain suggestions to problems that they admitted having, and the suggestions were arranged in order of their importance (i.e. suggestions to problems rated 5 were put first). Suggestions to problems that were rated "not at all" were not included. The second group (LI+PC) received the same personalised message as the first group did but they could not choose which person they would like to talk to. They were given

the lowly credible image used in Experiment 2 (photo 8). The third group (NI+PC) also received the same personalised message as the first group did but the message was shown without any image. The fourth group (NI+NPC) received a non-personalised message containing suggestions to all ten problems in a random order regardless of what they had selected. The message was shown with no image attached.

### Participants

Forty-four participants took part in the experiment. The participants' demographics are shown in Table 9 below. Participants were recruited online from an experiment listing website maintained by the University of Zurich (<http://genpsylab-wexlist.unizh.ch/>), and from a mailing list that staff and students of the university subscribed to. Participants came from all areas and professions. To guarantee that the participants replicate the actual potential users, we asked the participants whether they "currently want to exercise more regularly, but find it difficult" before the start of the experiment. Only participants who answered yes to this question were selected to complete the experiment.

	PI+PC		LI+PC		NI+PC		NI+NPC	
Number of participants	12		10		11		11	
Gender	M	F	M	F	M	F	M	F
	3	9	3	7	2	9	3	8
Average age	27.92 (13.70)		29.30 (7.63)		38.09 (12.05)		32.00 (11.95)	
	Y	N	Y	N	Y	N	Y	N
Exercising <sup>1</sup>	9	3	4	6	5	6	6	5
	5.08 (1.16)		4.80 (1.62)		5.45 (1.21)		4.18 (0.98)	
Knowledge <sup>3</sup>	5.50 (0.90)		5.60 (1.26)		5.82 (1.54)		5.27 (1.62)	
	4.83 (1.19)		4.70 (1.34)		4.36 (1.96)		3.82 (1.94)	
No of items read	6.42 (2.57)		7.60 (2.91)		6.36 (2.77)		10 (0.00)	

<sup>1</sup> I am currently doing some form of exercise

<sup>2</sup> I am more educated about my health than most people (1 = strongly disagree 7 = strongly agree)

<sup>3</sup> I have full knowledge of the benefits (consequences) of regular exercise (or the lack of it) (1 = strongly disagree 7 = strongly agree)

<sup>4</sup> I already knew all the information presented (1 = nothing 7 = everything)

**Table 9.** The participants' demographics.

### Results

The data collected in this experiment was also used to verify hypothesis 2, which postulated a positive correlation between the credibility of the person on a topic and the credibility of the message on that topic. Table 10 shows the Pearson correlation test results. The results are stronger than those of Experiment 2. The credibility of the source positively correlated with the credibility of the message when a photo is shown ( $r=0.619$ ,  $R^2=0.383$ ,  $p<0.01$ ) as well as when no photo is present ( $r=0.732$ ,  $R^2=0.536$ ,  $p<0.01$ ). Thus, H2 is supported. Interestingly, when looking at each individual group, the correlation between the

source's and the message's credibility was not significant for the PI+PC group, and the participants' demographics (Table 9) showed that the majority of people in this group are currently exercising. This is highly similar to our findings in Experiment 2.

		Source			Overall Credibility
		Attractiveness	Trust	Expert	
Web page credibility	PI+PC	0.349	0.479	0.326	0.462
	LI+PC	0.426	0.218	0.541	0.641*
	<b>With Photo</b>	<b>0.424*</b>	<b>0.471*</b>	<b>0.475*</b>	<b>0.619**</b>
	NI+PC	n/a	0.542*	0.785**	0.716**
	NI+NPC	n/a	0.490	0.749**	0.758**
	<b>No Photo</b>	<b>n/a</b>	<b>0.513**</b>	<b>0.763**</b>	<b>0.732**</b>

\* Correlation is significant at the 0.05 level (1-tailed)

\*\* Correlation is significant at the 0.01 level (1-tailed)

**Table 10.** Pearson's correlation between the credibility of the source and the message.

Hypothesis 3 stated that the presence of a personalised photo increases and the presence of a lowly credible photo decreases the credibility of the message. Similarly to previous experiments, we averaged the results on each dimension to get an overall score of credibility. The results of the experiment, shown in Table 11 below, reflect our prediction perfectly. The perceived credibility of the message when accompanied by a personalised photo was the highest, followed by the non-personalised message and the personalised message with no photo, and at last the personalised message with a lowly credible photo. There were, however, two unexpected results.

	F	D	G	T/E	Cred
PI + PC	5.03 (0.89)	3.97 (0.12)	5.69 (0.82)	4.77 (0.59)	<b>4.87</b> <b>(0.48)</b>
LI + PC	4.48 (0.63)	3.23 (1.36)	5.07 (0.78)	3.98 (1.06)	<b>4.19</b> <b>(0.49)</b>
NI + PC	4.84 (0.97)	3.52 (1.44)	5.27 (1.03)	4.50 (1.18)	<b>4.53</b> <b>(0.85)</b>
NI + NPC	4.25 (1.10)	4.03 (1.24)	5.24 (1.59)	4.77 (0.93)	<b>4.58</b> <b>(0.83)</b>

\* F = Fairness, D = Depth, G = Goodwill, T/E = Trust/Expertise, Cred = Credibility

**Table 11.** Average score and standard deviation of the perception of message credibility of each group.

First, only the difference between the PI+PC and the LI+PC group was statistically significant (independent t-test,  $p<0.01$ ). So, the credibility of the website with an image selected by the user was not significantly higher than that of the website without any image at all. There are two plausible explanations for this unexpected result. Firstly, participants may not have regarded the person in the chosen picture as the source of the message. We assumed that asking participants to choose someone whom they would like to talk to prior to presenting them a message with the photo of the selected person would give the illusion that the picture is the source of the message. However, choosing a picture and then finding it attached to the message is perhaps not the same as naturally finding a credible image already attached

to the message. The fact that subjects had to select from a large set of images (16) may well have made it less likely that they would assume there would be so many different contents.

A second possible explanation for the unexpected result is that participants may not have selected the most credible person for them. Indeed, while the average credibility of the person in the user-selected photo was significantly higher than that of the lowly credible photo (average=3.56, stdev=0.38 vs. average=3.29, stdev=0.41,  $p=0.05$ ), it was surprisingly not higher than that of the source when no photo was included (average=3.56, stdev=0.38 vs. average=3.76, stdev=0.70 (personalised message), and average=3.72, stdev=0.66 (standard message)). Only three out of twelve participants chose photo 12 or 10, which were the most favourite people to learn from about fitness programmes in Experiment 1 (see Table 4). Other chosen photos were: photo 6 (3 participants), photo 16 (1), photo 5 (3), photo 3 (1), and photo 4 (1). In addition, an average credibility of 3.56 for the person in the user-selected photo was considerably lower than what we would have expected considering the most credible person (photo 12) in Experiment 1 had a credibility of 3.87. Perhaps participants did not take the task of selecting a picture very seriously, knowing this was only an experiment, and suspecting that the person chosen would not influence the message.

These issues could be overcome by a more effective and unobtrusive way for identifying a user's preference for the appearance of the source. We will also consider reducing the number of images the user selects from, and more explicitly mentioning that the text will change depending on the person selected.

A second unexpected result is that our personalised content did not result in any improvement of credibility over the standard content. With hindsight, this is most likely due to the way personalisation is done: the standard content was longer (showing all ten information items) than the personalised content which only showed a subset (on average about three items less). Longer content is likely to result in a better rating for Depth (and there is a trend in that direction when comparing the NI+PC group with the NI+NPC group). Actually, studies in the area of recommender systems have shown that users appreciate longer reviews more [e.g. 18]. A more controlled follow-up study is needed to see if personalising the selection of and order in which the information items are presented has an effect when content length remains the same.

## 4 DISCUSSION

In this paper, we systematically investigated the effect of photographs of people on the perceived credibility of (the information) of a website.

### Limitations

To avoid any overgeneralization of our conclusions, we will first discuss the limitations of the study. Firstly, the majority of our participants are female, and over 30 years old. Hence, any findings in this paper should only be generalized to other demographics with care. Secondly, our findings are restricted to the subject of supporting people to exercise more regularly and planning a fitness programme.

## Findings and Implications

### (1) The credibility of a photo of a person is topic dependent.

Firstly, Experiment 1 showed that *the perceived credibility of a person as resulting from his/her appearance is topic dependent* (H1). The same person is likely to have different credibility when conveying information about different topics. This finding might explain the inconsistent results of a number of studies (e.g. [7, 9]), in which a photograph was assumed to have the same credibility when being used on different sites specialising in different products (digital cameras, computer hardware, and flower services in [7], and digital cameras, vitamins, and tour packages in [9]).

### (2) The more credible the person in the photo, the more credible the website. Adding a lowly credible photo can have a negative effect.

Secondly, we found a positive correlation between the credibility of the person in a photograph and the credibility of the website, and it accounts for some 30% of the variability in the website's credibility. So while the credibility of the person in the photograph is not the only factor that can influence the perception of the website's credibility, it clearly has an undeniable role. So while *adding the right image can positively affect the perception of trust, adding a lowly credible image can have a negative effect on the perceived credibility of the website*. These findings perhaps can further explain the mixed results found in other experiments where the effects of other variables might have not been accounted for, especially when the websites were more complicated.

### (3) Choosing the right photo is crucial yet challenging

It is trivial to see from our findings the importance and difficulty of choosing the right photo. But how can we do this?

Further qualitative analysis of Experiment 1 revealed that the participants apply different criteria in choosing someone who they would like the most to talk to about a certain topic. The most mentioned ones were "fit (but not overly fit)" (16/51 people), "expert" (16/51), friendly (6/51). These criteria are not only different between participants but also between the topic of the health benefits of exercise and planning personal fitness programmes (see [10, Section 2] for a more complete analysis). In fact, only 7 out of 51 participants chose the same favourite person for the two topics. In a series of studies by Baylor [13] where participants selected a human-like agent to act as their teacher, participants had the tendency to select an agent that was highly similar to themselves in terms of gender and ethnicity. These user-selected agents were perceived to be more credible and motivating than a standard agent or no agent. An interview of thirty Internet shoppers by Keeling [14] also revealed that what is considered the appropriate appearance for an agent attached to an e-commerce website not only depends on the users but also the product sold on the site. This suggests that *finding a generic image that can be liked and perceived as highly credible by the majority of users would be considerably challenging, especially when there are great varieties in the user population*.

There is a clear need for an effective and unobtrusive method for identifying a user's preferences for the appearance of onscreen characters. Our initial attempt of simply asking the

users to select someone whom they would like to speak to about a certain topic did not work as well as expected. On a positive note, adding a user-selected photograph yielded a significant improvement on both the perceived credibility of the source and that of the information over adding a lowly credible photograph with respect to the topic of the website. However, although adding a user-selected photograph enhanced the credibility of the website when no photo was present, the difference was not significant. Was it because the participants did not choose the most credible photo in the experimental condition? Was it because selecting a photo and then finding it attached to a message does not lead to the perception that the chosen photo is the source of the message? The responses from the participants in Keeling's study [14] showed that the participants indeed "varied in the extent to which they regarded the ECA as presenting the information". These issues will be addressed in a future study.

### Conclusions

In summary, our findings clearly influence the process of using photographs (and onscreen characters in a broader sense) to improve the credibility of websites. While adding such characters can have a positive impact, choosing the right appearance is crucial, as the wrong appearance can lead to the opposite of the intended effect.

### Future Work

Firstly, the vast majority of studies regarding the relationship between displaying photographs of people on websites and the sites' credibility, including our own, do not investigate how this relationship develops over time. Adding a photograph might be able to create a good first impression, but will it last? How will it increase or decrease over multiple interactions? We plan to explore this question further.

Secondly, we plan to compare how different forms of a source's visual appearance (e.g. animated characters, photographs of people, videos of people, voice, etc.) affect the credibility of a website.

Thirdly, we plan to further investigate the tailoring of images to users. As mentioned above, our first attempt of letting users select their favourite person to talk to was not very successful. We will investigate whether we can obtain better results by allowing users to select a person from a smaller set and more explicitly telling users that the content will change depending on the person they select. We will also investigate whether automatically selecting an image based on user stereotypes has a positive effect.

Finally, the use of onscreen characters is not restricted to having one character interacting and delivering the information to the users in a direct communication setting. For instance, multiple characters can be used to simulate the positive effect of social norms, and indirect communication on the perception of the credibility of the information [1,2]. Such impact is worth investigating, yet there have been few studies addressing it [15-17].

### ACKNOWLEDGEMENTS

Judith Masthoff's research is partly supported by Nuffield

Foundation Grant No. NAL/00258/G. We would like to thank the anonymous reviewers for their improvement suggestions.

### REFERENCES

- [1] J. B. Stiff. *Persuasive Communication*. The Guildford Press, New York, (1994).
- [2] E. Aronson. *The Social Animal*. Worth, (2004).
- [3] S. Parise, S. Kiesler, L. Sproull, and K. Waters. Cooperating with Life-like Interface Agents. *Computers in Human Behavior*, Vol 15 No 2, 123-142 (1999).
- [4] B. J. Fogg, J. Marshall, T. Kameda, J. Solomon, A. Rangnekar, J. Boyd, and B. Brown. Web Credibility Research: A Method for Online Experiments and Early Study Results. *CHI 2001*, 295-296 (2001).
- [5] P. De Vries. Social Presence As A Conduit to The Social Dimensions of Online Trust. *Persuasive Technology Conference*. Eindhoven, The Netherlands, 55-59 (2006).
- [6] U. Steinbrueck, H. Schaumburg, S. Duda, and T. Krueger. A Picture Says More Than A Thousand Words – Photographs As Trust Builders In E-Commerce Websites. In *Proceedings of CHI 2002: Extended Abstracts*. Minneapolis, MN, US, 748-749 (2002).
- [7] J. Riegelsberger, M. A. Sasse, and J. D. McCarthy. Shiny Happy People Building Trust? Photos on E-commerce Websites and Consumer Trust. *Proceedings of CHI 2003*, 121-128 (2003).
- [8] J. Riegelsberger, and M. A. Sasse. Face It: Photographs Don't Make Websites Trustworthy. In *Proceedings of CHI2002: Extended Abstract*, Minneapolis, MN, US, 742-743 (2002).
- [9] P. De Vries, and A. Th. H. Pruyn. Source Salience and The Persuasiveness of Peer Recommendations: The Mediating Role of Social Trust. In *Proceedings of the 2<sup>nd</sup> Persuasive Technology Conference*. California, CA, US, Springer Verlag, 164-175 (2007).
- [10] H. Nguyen, and J. Masthoff. Is It Me or Is It What I Say? Source Image and Persuasion. In *Proceedings of the 2<sup>nd</sup> Persuasive Technology Conference*. California, CA, US, Springer Verlag, 231-242. (2007).
- [11] R. Ohanian. Construction and Validation of A Scale To Measure Celebrity Endorsers' Perceived Expertise, Trustworthiness, and Attractiveness. *Journal of Advertising*, Vol 19 No 3, 39-52 (1990).
- [12] T. Hong. Contributing Factors To The Use of Health-related Websites. *Journal of Health Communication*, Vol 11, 149-165 (2006).
- [13] A. L. Baylor, R. B. Rosenberg-Kima, and E. A. Plant. Interface Agents as Social Models: The Impact of Appearance on Females' Attitude Toward Engineering. *CHI 2006*, Montreal, Canada (2006).
- [14] K. Keeling, S. Beatty, P. McGoldrick, L. Macaulay. Face Value? Customer Views of Appropriate Formats for Embodied Conversational Agents (ECAs) in Online Retailing. *Proceedings of the 37<sup>th</sup> Hawaii International Conference on System Sciences*, (2004).
- [15] S. D. Craig, B. Gholson, M. H. Garzon, X. Hu, W. Marks, P. Wiemer-Hastings, Z. Lu. Autor Tutor and Otto Tudor. *AIED-Workshop on Animated and Personified Pedagogical Agents*, Le Mans, France, 25-30 (1999).
- [16] E. Andre, T. Rist, S. van Mulken, M. Klesen, S. Baldes. The Automated Design of Believable Dialogues for Animated Presentation Teams. In *J. Cassell, S. Prevost, J. Sullivan, and E. Churchill: Embodied Conversational Agents*, The MIT Press, 220-255 (2000).
- [17] H. Nguyen, J. Masthoff, P. Edwards. Persuasive Effects of Embodied Conversational Agent Teams. *Proceedings of 12th International Conference on Human-Computer Interaction*, Beijing, China, Springer-Verlag, Berlin, 176-185 (2007).
- [18] Tintarev, N. and Masthoff, J. (2007). Effective explanations of recommendations: User-centered design. *Proceedings of the ACM Conference on Recommender Systems*, Minneapolis, US, 153-156.