

Designing Robot Embodiments for Social Interaction: Affordances Topple Realism and Aesthetics

Robert A. Paauwe¹ · Johan F. Hoorn¹ · Elly A. Konijn¹ · David V. Keyson²

Accepted: 5 May 2015 / Published online: 7 June 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract In the near future, human-like social robots will become indispensable for providing support in various social tasks, in particular for healthcare (e.g., assistance, coaching). The perception of realism, in particular human-like features, can help facilitate mediated social interaction. The current study investigated the effects of form realism on engagement with and use intentions of social robot embodiments. We have defined (perceived) form realism as the result of the appraisal of features that are perceived as realistic contrasted with those appraised as unrealistic. To test the effects of form realism, we applied the model of interactively perceiving and experiencing fictional characters (I-PEFiC). I-PEFiC explains how users respond to interactive, fictional, humanoid characters, on social robots. In a within-subjects design, participants ($N = 29$; $M_{age} = 28.8$ years, age range 18–56 years) interacted with three different robots built from LEGO Mindstorms, which differed in their degree of designed form realism. Each robot presented itself as a physiotherapy assistant and requested the participant to do several exercises. Results of a structured questionnaire indicated that form realism only played a modest role in the perception of electro-mechanical robots. Instead, the perception of affordances appeared to be crucial for determining engagement and intentions to use social robots.

Keywords Social robots · Realism · Engagement · Use intentions · Design

✉ Robert A. Paauwe
r.a.paauwe@vu.nl

¹ Faculty of Social Sciences, VU University Amsterdam,
De Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

² Faculty of Industrial Design Engineering,
Delft University of Technology, Landbergstraat 15,
2628 CE Delft, The Netherlands

1 Introduction

Globally, estimates are that over 1 billion people aged 65 and older will live in 2030, compared to ‘only’ 500 million in 2006 [9]. More recent figures estimate that this population will grow beyond 1.5 billion in 2050 [37]. The increase in life expectancy, together with an increase in the total amount of required healthcare will lead to a high demand of care professionals. Furthermore, the amount of available healthcare professionals (e.g., caregivers, nurses) is expected to decline in the near future due to a lack of (qualified) employees and resources [27,36]. Therefore, healthcare services become more expensive and services become increasingly difficult to facilitate by humans. It is likely that at some point social robots will become indispensable in supporting care professionals or that they will directly take care of patients (e.g., [6,7]). The present paper aims to contribute to the developing role of robots in the care sector by providing scientific building blocks for the design of social robots.

1.1 Realism

To help ensure the quality of healthcare services in terms of ensuring human-like interaction, social robots should be able to express their own engagement and emotions, show empathy, and have the ability to interpret the human equivalents in the interaction partner [10,34]. Likewise, realism is an important factor for eliciting social cues tied to these aspects [1–3] when interacting with virtual agents and robots. Thus far however, media effects research has shown mixed results for the role of perceived realism, which is partly due to how it is defined [5,46]. Realism manifests itself in form, behavior, and visual fidelity [2,3,5,40,46].

1.1.1 Form Realism

Form realism relates to the degree of realism of the physical appearance of the robot (i.e. embodiment) [2,3]. We define perceived form realism as the result of subjective appraisals of features as realistic contrasted with those features that are designed in the stimulus and appraised as unrealistic (cf. [15, p.62] and [33]). Form realism also appears to be similar to representation or resemblance, which refers to the extent an embodied agent resembles its corresponding real world entity (e.g., anthropomorphism, zoomorphism [2,11]). Anthropomorphism relates to how much an entity is similar to a real human. By humanizing the appearance of social robots, engagement might be increased and allows users to employ social and physical conventions that are familiar to them (cf. [20,31]). Furthermore, increasing anthropomorphism leads to longer interactions with the robot and a more positive, lifelike evaluation of the robot [19]. A human-like agent also triggers stronger social reactions compared to completely textual agents [1,35]. Human appearance appears to be important for social interaction, yet Nowak and Biocca demonstrated that participants were engaged with a virtual agent, regardless whether the avatar looked human or not [28]. Furthermore, a robot that resembles a human too much could make its social shortcomings evident and too few human-like features might lead to underestimating the communicative skills of the robot [8,10,11].

1.1.2 Visual Fidelity

Visual fidelity (also sometimes defined as photographic realism or photorealism) refers to the degree of realism in the used medium [25,46]. For example, the same virtual agent can be rendered using computer graphics as photographic or as cartoon-like. Where form realism concerns what the embodiment represents, visual fidelity concerns the visual quality of the representation. Yee et al. demonstrated in a meta-analysis that human-like representations with high visual fidelity evoked more positive social interactions than those with low visual fidelity [46]. However, a too realistic depiction of a human being also may result in an uncanny or eerie experience (cf. [26]). This unsettling experience is likely due to a mismatch between appearance and behavior. At higher degrees visual fidelity in human faces, unpleasant experiences emerge, in particular related to atypical (abnormal) features [32]. Furthermore, it appears that visual fidelity does not affect the social influence an artificial entity has [44].

1.1.3 Behavioral Realism

Behavioral realism concerns the demeanor of the robot (e.g., expression, movement, gestures, gaze [2,3,34,40]). It is the degree to which the behavior of the embodiment

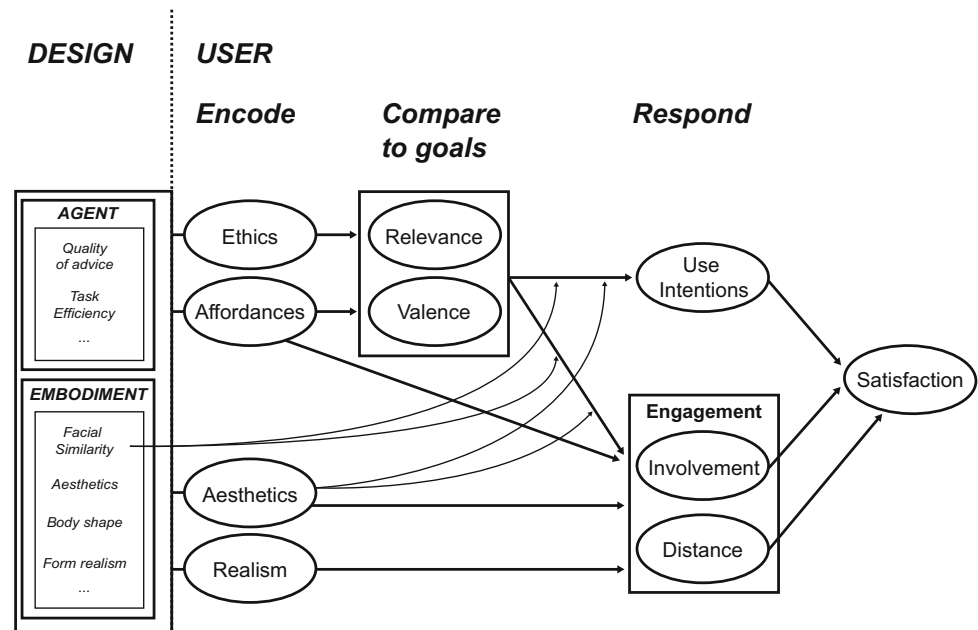
matches with what it is supposed to resemble (e.g., ‘does it behave realistic according to what it represents?’). This match between behavior and representation is determined by many different aspects (e.g., movement, interactive behavior, social behavior, verbal, and non-verbal communication [41,44]). Higher behavioral realism leads to increased perceived social cues and evokes more social responses in the user [30]. However, if there is a mismatch between behavioral realism and form realism, social presence declines [2,3]. Therefore, it appears to be key to harmonize these parameters of behavioral realism (cf. [41]) and match behavioral realism with form realism and visual fidelity.

1.2 Physical Embodiment

One of the differences between digital interface agents and electro-mechanical robots is that the latter are physically present in 3D space, whereas interface agents are present on a screen. Lee [23, p. 37] defines presence as “a psychological state in which virtual (para-authentic or artificial) objects are experienced as actual objects in either sensory or non-sensory ways”, which can occur at the physical level, the social level, and the personal (self) level. Interacting with a robot also appears to be more compelling and engaging compared to interacting with an avatar [19]. Several other studies [17,24] indicated that robots with a physical embodiment have a positive effect on feeling the social presence of the robot, as compared to identical robots with a digital embodiment. Both studies concluded that having a physical embodiment contributes to experiencing social presence of robots. [17,24]. However, if tactile interaction with the robot is limited, having a physical embodiment could have negative effects on experiencing social presence [17,24]. Several pilot studies [42,43] hinted at the positive contribution of physical embodiment to the perceived watchfulness, perceived helpfulness, and enjoyability of a social robot.

Kidd and Breazeal demonstrated that physical presence is not necessarily important for robot engagement, but rather the perception that something exists and is experienced as real [18]. In their study, users perceived physical robots (a physically present robot and a televised robot) as something that exists somewhere in the real world, but a digitally animated avatar as something fictional. If an embodiment is perceived as something that exists in the real world (even when televised), engagement, credibility, informativeness, intelligence, and enjoyment are all positively affected [4,18]. Additionally, humans employ more social attention when they believe if a (robot) embodiment is controlled by a human compared to being controlled by a machine, regardless of representation [45]. However, for eliciting positive social interactions, having some form of representation is better than not having any representation [46].

Fig. 1 Interactively perceiving and experiencing fictional characters. [38]



The belief that something exists and resembles something ‘real’ is related to the design of the robot embodiment. What the robot represents may be of influence in whether users believe a robot is real, and therefore more engaging and likely to be used. By focusing on form realism (i.e. what the robot embodiment resembles), we can determine whether the embodiment of the robot (regardless of behavior or visual quality) affects the user’s emotional responses or their intention to interact with the robot. We propose the following hypothesis:

H1 A social robot with a greater degree of form realism designed in its embodiment arouses a higher level of perceived form realism in the user.

1.3 The I-PEFiC Framework

To investigate the effects of designed form realism of physical robot embodiments on the user, we applied the model of I-PEFiC, (see Fig. 1). I-PEFiC explains how users interact with fictional characters based on the user’s perception of ethics, affordances, aesthetics, and realism [38]. I-PEFiC divides the process of perceiving and experiencing interactive (fictional) characters into three separate phases: encode, compare, and respond (see Fig. 1).

When interacting with fictional characters (e.g., avatars), humans perceive features of these characters in terms of ethics (how good and evil is the character?), affordances (is the character aiding or obstructing the users goals?), aesthetics (how beautiful and ugly is the character?) and realism (how real and fake is the character?). After encoding, users compare these features to their own characteristics, goals, concerns, and respond accordingly [22,38], resulting in a felt

tendency to use the character or not (use intentions). This is accompanied by a level of engagement with the character. In I-PEFiC, engagement comprises the parallel experience of involvement and distance towards a fictional character [21,22]. Involvement can be explained as the inclination to affectively approach the character someone is interacting with (e.g., feeling sympathy), and distance is the experienced tendency of wanting to avoid the character (e.g., being annoyed) [22]. Finally, the combination of use intentions together with involvement and distance explain the overall satisfaction with the character [38–40].

We do realize that (product) affordances usually are defined as the potential action possibilities that a user perceives with a product, in relation to a certain context (cf. ecological affordances, [12]). However, this is difficult to operationalize in the form of a questionnaire and introduces additional test fatigue in the participants, due to the extensive set of factors in I-PEFiC. Therefore, similar to previous I-PEFiC studies [21,22,38–40], we have aggregated affordances into an overall level of competence (e.g., ‘is the robot capable to successfully help you?’), rather than the actual amount of perceived action possibilities.

Previous studies empirically validated PEFiC (at the time without interactivity) in relation to media characters [21,22] and I-PEFiC with interface agents [39,40]. However, no studies were conducted with physically embodied robots. Using I-PEFiC as our theoretical framework, we designed an experiment to investigate the effect of designed form realism of social robots on the engagement and use intentions. We specifically focused on I-PEFiC’s tangible design factors: affordances, aesthetics, and realism, and their relation with the interaction and engagement processes. The factors ethics, relevance, and valence were left out because they are

less related to (industrial) design. Furthermore, it is important to examine whether I-PEFiC can serve as an appropriate model for predicting engagement with and intentions to use physically embodied social robots. Based on prior studies, we hypothesized the same interrelationships between the factors in the I-PEFiC model as established in previous research (cf. Fig. 1). We propose the following hypotheses:

H2 Perceived affordances have a direct, positive effect on use intentions.

H3 The relationship between perceived affordances and use intentions is moderated by perceived aesthetics.

H4 Perceived affordances have (a), a direct positive effect on involvement, as well as (b), a direct negative effect on distance.

H5 Perceived aesthetics has (a), a direct positive effect on involvement, as well as (b), a direct negative effect on distance.

H6 Perceived realism has (a), a direct positive effect on involvement, as well as (b), a direct negative effect on distance.

2 Method

2.1 Participants and Study Design

The study included 29 participants (13 female, 16 male, $M_{age} = 28.8$ years, age range 18–56 years) in a within-subjects design. Each participant interacted with three robots that varied in what they resembled (human \times creature \times machine). Twenty-seven adults were recruited from a group

of university level adults (M.Sc. and Ph.D. students) and two additional participants were recruited from a filing department (university employees). The participants were recruited from the faculty of social sciences and none of them had prior experience with robots. Furthermore, the participants did not receive any compensation for taking part in the experiment.

2.2 Materials

Three robot designs that differed in their form realism (e.g., what they resembled) were created, resulting in a human-like robot ‘Rob’, a crocodile-like robot ‘Croq’, and a vehicle-like robot ‘Car’. We suspected that varying within a single embodiment design (e.g. three human-like robots) would lead to too subtle responses. Therefore, three different morphologies were chosen in an attempt to maximize the effects they might have on I-PEFiC factors. Rob, Croq, and Car were created from three LEGO Mindstorms NXT 2.0 sets (set #8527), which are toy sets intended for building small, programmable robots. The LEGO Mindstorms enabled us to use the same bricks such that the basic construction elements were similar across the three types. Only the exterior form (different representations) of the robots varied, whilst keeping visual fidelity and aesthetics consistent for each of the three robots. To maintain consistency in behavioral realism as well, all of the robots were static; none of them moved or was animated. The designs of Rob, Croq, and Car were based on the design templates included in the LEGO Mindstorms NXT 2.0 kit and roughly used the same number of LEGO bricks. The ‘Alpha Rex’ design served as a base for Rob, the ‘Robogator’ design for Croq, and the ‘Shooterbot’ design for Car. Each robot design was modified such that they included an ultrasonic sensor, a microphone, a color sensor, and a push button (see Fig. 2, left image).



Fig. 2 The three social robots used in the experiment; Rob, Croq, and Car (from left to right in the left image) and the experimental setup (right image). The adjacent rooms for Croq and Car were similar to the room of Rob

Table 1 Overview of the used measures

Measure	Items*	Cronbach's α
Resemblance	...looks like a real human, ...looks like a real crocodile, ...looks like a real car	**
Perceived affordances	skilled, proficient, clever, competent, <i>clumsy</i> , dumb, lacking, inapt	0.88
Perceived aesthetics	pleasant to look at, pretty, attractive, <i>unpleasant to look at</i> ugly, <i>misshapen</i>	0.72
Perceived realism	real, natural, fake, artificial	0.81
Use intentions	use it again, follow its advice, visit it again, ignore it, remove it, get rid of it, continue working with it, do another task with it, <i>complete the task on my own</i> , prefer to work alone	0.97
Involvement	good, <i>attached</i> , pleasant, comfortable, happy, joyful, satisfied	0.92
Distance	annoyed, awful, uncomfortable, <i>sad</i> , <i>detached</i> , irritated, unpleasant	0.91

* Items in italic are discarded based on the results of the PCAs and are not included in the reliability analysis

** Scale reliability was not calculated for the manipulation check

Rob, Croq, and Car were controlled by identical software and were programmed in C#, using Microsoft Visual Studio 2010 to execute the program. The speech of the robots was synthesized using Loquendo Text To Speech (v7.20.2), with each robot having the same, neutral male voice. The language spoken by the robots was Dutch and was projected via speakers hidden directly behind each robot. The robot provided instructions for four different morning gymnastics exercise. Rob, Croq, and Car were designed as a physiotherapy robots because gymnastics exercises are a type of healthcare interventions that are easy to relate to for participants that are not patients or in direct need of care. The participant activated the exercise routines by presenting colored (red, green, blue, yellow) cardboard cards (85 mm \times 54 mm) to the color sensor of the robot (see Sect. 2.3).

2.3 Procedure

There was a single session leader overseeing the experiment and was waiting in the hallway during the experiment. Rob, Croq, and Car were placed in similar, adjacent rooms connected to the hallway where the session leader was positioned. Each door was marked with a sign displaying the name of the robot that was present in that specific room. Because every participant had to interact with all three robots, each participant received an instruction envelope from the session leader before the experiment started. The envelope contained an instruction card that stated in which order they had to interact with Rob, Croq, and Car. The orders in which participants had to interact with the three robots were evenly and randomly distributed among the participants.

With this setup, three participants could do the experiment simultaneously with only one participant at the same time in a room interacting with a robot. Participants were given specific time slots to participate in the experiment; no participants were waiting at the hall for their turn while the experiment was in progress. Furthermore, the rooms were

blinded so participants could not look inside the room from the hallway. Once the participants entered the appropriate room, the session leader informed them that they could start by pressing the button next to the robot (see Fig. 2, right image). The session leader left the room after giving the instructions, returning to the hallway.

The robot introduced itself as Rob / Croq / Car, the physiotherapy robot, asking the participant to present one of four colored cards to its sensors, each card corresponded with a different gymnastics exercise routine. To reduce potential test fatigue, the participants were given the freedom to determine the sequence in which they presented the cards to the robots. The red card started an exercise to clap as loudly as possible with the arms stretched out. Once completed, the robot asked the participant to clap again, but louder than the first time. The green card started a hand-eye coordination exercise, in which the robot asked the participant to close their left eye and press the button with their right index finger. When the task was completed, the robot asked the participant to press the button again, now with their right eye closed using their left index finger. The blue card started a stretching exercise in which the robot asked the participant to stretch their right arm towards the face of the robot, keeping it stretched. When the participant correctly stretched their right arm, the robot asked to do the same exercise again, now using the left arm. Finally, the yellow card started another stretching exercise in which the robot asked the participant to hold both fists at the chest and stretch them both together towards the face of the robot. The robot counted every stretch out loud and after stretching five times, the exercise was completed. Once all cards were presented and the exercises were completed successfully, the robot thanked the participant and requested to fill out the paper questionnaire (see Table 1 and Sect. 2.4 for an overview of the measures used and reliabilities).

The robots were entirely autonomous; they detected colored cards with a color sensor, correct movements with an ultrasonic sensor, sounds with a microphone, and pressure by

means of a push button. Once the tasks and the questionnaire were completed, the participants were instructed to step outside of the room and wait until all of the participants within their group were finished and then to continue with the next robot as indicated on their instruction card. When the participants had interacted with all three robots and completed the questionnaires accordingly, they were debriefed by the session leader and thanked for their participation.

2.4 Measures

Each questionnaire was divided into five blocks of items, and the order in which the blocks were presented, as well as individual items within the blocks, were pseudo-randomized using an algorithm. No more than two items from the same variable were presented in a row (e.g., aesthetics item 1, aesthetics item 2, affordance item 1, etcetera). The questionnaire in the first condition of each participant included demographic questions. Items were taken from previous I-PEFiC studies [39,40] and translated for the purpose of the present study. Each factor was measured with 4 to 10 Likert-type items, as described below. Each item was followed by 6-point rating scales, ranging from 0 = ‘totally disagree’ to 5 = ‘totally agree’. Items were positively and negatively formulated; the negative items were recoded after the experiment for composing the scales. To get an indication whether the items were coherent within their corresponding scales and distinguishable from the others, principle component analyses (PCAs) were done. Missing data was imputed using multiple imputation (random seed = 20070525). Because less than 10 % of the data was missing, 5 additional datasets were imputed ($\gamma > 0.1$, $m = 5$, see [13]). The pooled dataset was used if the analysis allowed it, otherwise the first imputed dataset was used. All resulting scales were reliable according to Cronbach’s α (> 0.72).

2.4.1 Resemblance

The manipulation check items measured the resemblance of the robots to their corresponding real-world entities, targeting the resemblance of Rob, Croq, and Car individually. Resemblance was measured with 2 items starting with the phrase: ‘Rob / Croq / Car looks...’ and were completed by ‘...like a real human / crocodile / vehicle’ respectively and ‘...like a machine for all’.

2.4.2 Perceived Affordances

The Perceived Affordances of Rob, Croq, and Car were measured with 8 items, starting with the opening phrase: ‘For a physiotherapy robot, I think Rob / Croq / Car is...’, followed by 4 items that expressed positive affordances (skilled, proficient, clever, competent) and 4 negative ones (clumsy, dumb,

lacking, inapt). Scale reliability (Cronbach’s $\alpha = 0.88$) was calculated after discarding one item (clumsy) due to ambiguity in the PCA.

2.4.3 Perceived Aesthetics

The Perceived Aesthetics of the robots was measured with six items and started with the phrase: ‘For a physiotherapy robot, I think Rob / Croq / Car is...’, followed by 3 positive statements on Aesthetics (pleasant to look at, pretty, attractive) and 3 negative ones (unpleasant to look at, ugly, misshapen). Out of these 6 items, 2 items (unpleasant to look at, misshapen) were discarded based on PCA because they lacked coherence with the other items. Scale reliability was calculated after discarding these two items (Cronbach’s $\alpha = 0.72$).

2.4.4 Perceived Realism

Perceived Realism of Rob, Croq, and Car was measured with 4 items, starting with the phrase: ‘Rob / Croq / Car looks...’, followed by 2 positively formulated items (real, natural) and 2 negatively formulated items (fake, artificial). Scale reliability was calculated (Cronbach’s $\alpha = 0.81$). None of the items were discarded based on the results of the PCA.

2.4.5 Use Intentions

The intention to use Rob, Croq, and Car was measured with 10 items, 5 of which were expressed positively and 5 negatively. The items were divided into one block of 6 items, and one block of 4 items. The first block started with the phrase: ‘I would like to...’ and was completed by 6 items (use it again, follow its advice, visit it again, ignore it, remove it, get rid of it). The second block used the initial phrase: ‘The next time, I would...’ and was concluded by 4 items (continue working with it, do another task with it, complete the task on my own, prefer to work alone). One item of Use Intentions (complete the task on my own) was discarded based on PCA indication of incoherence with the other nine items. Scale reliability was calculated after discarding this item (Cronbach’s $\alpha = 0.97$).

2.4.6 Involvement

Involvement with the robots was measured with 7 items starting with: ‘When dealing with Rob / Croq / Car, I felt...’ followed by seven expressions (good, attached, pleasant, comfortable, happy, joyful, satisfied). PCA revealed that one item (attached) had no coherence with any of the other items. After discarding this item, scale reliability for Involvement was calculated (Cronbach’s $\alpha = 0.92$).

2.4.7 Distance

Distance towards the robots was measured with 7 items, opening with: ‘When dealing with Rob / Croq / Car, I felt...’ and followed by distance-related expressions (annoyed, awful, uncomfortable, sad, detached, irritated, and unpleasant). Two items (sad, detached) were discarded as a result of the PCA due to incoherence of these items compared to the other five. Scale reliability was calculated (Cronbach’s $\alpha = 0.91$).

3 Results

3.1 Manipulation Check

To check whether our manipulation of designed form realism was successful, we tested how participants perceived the realism designed in Rob, Croq, and Car in terms of Resemblance to their reference entities in the real world (cf. [2]). That is, we analyzed to what degree Rob resembled a human being, Croq a real crocodile, and Car a car. A one sample *t*-test with 0 as the test value revealed that the Resemblance rating of Rob was significantly higher than 0 on a 6-point rating scale ($M = 1.28, SE = 1.98, t(28) = 6.45, p < 0.001$). The Resemblance rating of Croq was also significantly higher than 0 ($M = 2.00, SE = 0.258, t(28) = 7.76, p < 0.001$). Finally, the Resemblance rating of Car was significantly higher than 0 ($M = 3.10, SE = 0.269, t(28) = 11.54, p < 0.001$). In other words, Car was perceived with the highest level of resemblance to the real-life object, followed by Croq and Rob respectively. Thus, our manipulation of designed realism can be considered successful in terms of sufficiently looking like the real-world entity.

Additionally, we looked whether the three representations differed among one another. We performed a one-way repeated measures MANOVA with Designed Form Realism and Resemblance. At the multivariate level, the results revealed a significant main effect of Designed Form Realism

(Wilks’ $\lambda = 0.40, F(4, 25) = 9.59, p < 0.001, \eta_p^2 = 0.61$). Univariate results confirmed that Designed Form Realism had a significant effect on the level of Resemblance ($F(2, 56) = 23.36, p < 0.001, \eta_p^2 = 0.46$). Contrasts indicated that the difference between Rob and Car ($F(1, 28) = 23.36, p < 0.001, \eta_p^2 = 0.57$) was the largest, followed by Croq and Car ($F(1, 28) = 35.31, p < 0.01, \eta_p^2 = 0.39$), and finally Rob and Croq ($F(1, 28) = 8.55, p < 0.01, \eta_p^2 = 0.23$). Thus, our manipulation of designed realism was successful in that all three representations differed from each other in their degree to which they resembled their corresponding real-world entity.

3.2 Effect of Designed Form Realism on Perceived Realism

To test whether a social robot with a greater degree of Designed Form Realism (embodiment) would lead to a higher degree of Perceived Form Realism (H1), we conducted a one-way repeated measures MANOVA. Multivariate results revealed no main effect of Designed Form Realism on the levels of perceived realism across the three designs of Rob, Croq, and Car (Wilks’ $\lambda = 0.94, F(6, 22) = 2.44, p = 0.96, \eta_p^2 = 0.06$). Therefore, H1 had to be rejected (see Table 2). Because no significant differences were found in the perceptions of realism as a function of designed form realism, we collapsed the scores of Rob, Croq, and Car for the analysis of H2 through H6 (see Table 2), thereby increasing the statistical power.

3.3 Perceived Aesthetics × Perceived Affordances on Use Intentions

To test the assumed interaction of Perceived Aesthetics and Perceived Affordances on Use Intentions (H2 and H3), we used the SPSS macro PROCESS [14], which is a macro using a regression-based approach to analyze mediation and moderation models together with resampling techniques

Table 2 Overview of linear regression analyses for H2 through H6

Effect of	<i>p</i>	<i>b</i> (<i>SE_b</i>)	β	CI
H2: Perceived affordances on use intentions	0.40	0.29(0.34)	*	−0.67–1.24
H3: Perceived affordances × perceived aesthetics on use intentions	0.04	0.31(0.15)	*	−0.10–0.71
H4a: Perceived affordances on involvement	0.001	0.44(0.11)	0.49	
H4b: Perceived affordances on distance	<0.001	−0.86(0.13)	−0.84	
H5a: Perceived aesthetics on involvement	0.21	0.20(0.16)	0.20	
H5b: Perceived aesthetics on distance	0.84	0.04(0.18)	0.03	
H6a: Perceived realism on involvement	0.03	0.38(0.17)	0.34	
H6b: Perceived realism on distance	0.99	0.01(0.19)	0.003	

* PROCESS reports confidence intervals rather than standardized coefficients (see [14])

(bootstrapping). To test the interaction, 5000 bootstrap samples were generated and tested the significance of the interaction effect at 99 % bias-corrected bootstrap confidence intervals, using moderation model 1 [14, p. 218 and 442]. The analysis revealed that the model in its entirety explained 78 % of the variance in Use Intentions ($F(3, 25) = 30.29, p < 0.001$). When looking at individual factors, only the interaction between Perceived Affordances and Perceived Aesthetics on Use Intentions was significant ($b(SE_b) = 0.31 (0.15), p = 0.04$). Perceived Affordances ($b(SE_b) = 0.29 (0.34), p = 0.40$) and Perceived Aesthetics ($b(SE_b) = -0.63 (0.42), p = 0.15$) did not relate to Use Intentions when this interaction was considered. Therefore, we concluded that Perceived Aesthetics moderated the effect of Perceived Affordances on Use Intentions, which is in support of H3, but not H2 (see Table 2).

3.4 Effect of Perceived Affordances, Perceived Aesthetics, and Perceived Realism on Involvement

To test H4a, H5a, and H6a, we performed a linear regression analysis with Perceived Affordances, Perceived Aesthetics, and Perceived Realism as predictor variables and Involvement as the dependent variable. The analysis revealed that these three Encode factors explained 67 % of the variance in Involvement ($F(1, 25) = 37.78, p < 0.001$). When looking at individual factors, the analysis revealed a relation between Involvement and Perceived Affordances ($b(SE_b) = 0.44 (0.11), \beta = 0.49, p = 0.001$) as well as Perceived Realism ($b(SE_b) = 0.38 (0.17), \beta = 0.34, p = 0.03$), but not with Perceived Aesthetics ($b(SE_b) = 0.20 (0.16), \beta = 0.20, p = 0.21$). These findings support H4a, and H6a, but not H5a (see Table 2).

3.5 Effect of Perceived Affordances, Perceived Aesthetics, and Perceived Realism on Distance

To test H4b, H5b, and H6b, we performed a linear regression analysis with Perceived Affordances, Perceived Aesthetics, and Perceived Realism as predictors, and Distance as the dependent variable. The analysis revealed that the model in its entirety explained 68 % of the variance in Distance ($F(1, 25) = 57.28, p < 0.001$). When looking at the individual factors, the analysis revealed a direct relationship between Distance and Perceived Affordances ($b(SE_b) = -0.86 (0.13), \beta = -0.84, p < 0.001$), but not with Perceived Aesthetics ($b(SE_b) = 0.04 (0.18), \beta = 0.03, p = 0.84$) or Perceived Realism ($b(SE_b) = 0.01 (0.19), \beta = 0.003, p = 0.99$). These findings support H4b, but not H5b nor H6b (see Table 2).

3.6 Exploratory Analyses

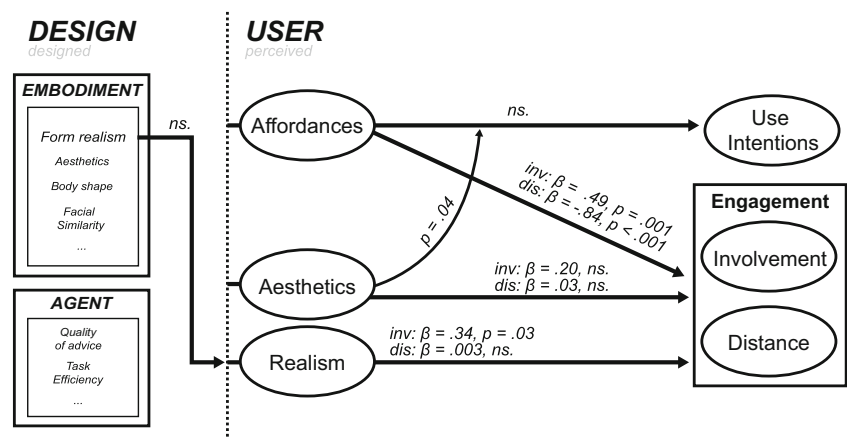
Finally, to explore whether age, gender, or prior robot experience affected the results of the main hypotheses, we performed a one-way repeated measures MANCOVA. Gender was treated as a between-subjects factor, and age, profession, and prior robot experience were included as covariates. No significant interaction effects were found between the manipulated Designed Form Realism and gender, age, profession, or prior robot experience ($ps > 0.05$).

4 Discussion

The goals of this study were to test the effects of Designed Form Realism on Use Intentions, Involvement, and Distance in the interaction between humans and social robots with physical embodiments, and to examine whether the I-PEFiC framework is applicable for this domain. Results showed that despite a successful manipulation of Designed Form Realism (i.e. differences in resemblance), participants did not perceive any differences between Rob, Croq, and Car in terms of realism (H1). That is, no significant differences showed up between the three variants of our robots; all three robots were perceived roughly similar on these dimensions. Presumably, Car's resemblance to its real-life object was the highest because a LEGO vehicle and a real vehicle are both mechanical objects, whereas the real-life representation of Croq and Rob (i.e., a crocodile and human, respectively) are not mechanical. Even though the participants perceived a resemblance between the individual robots and their corresponding real-world entities, as well as different degrees of resemblance between the robots, the degree to which they experienced realism of the robots was unaffected. Interestingly, perceived realism of social robots did not appear to be important for the interaction or engagement of participants with social robots (supporting H6a, not H6b). This finding is in accordance with previous studies [16, 39, 40, 46], concluding that realism only plays a modest or sometimes no role in the perception of virtual partners.

The perception of affordances, (here, gymnastics), appeared to be much more important for user engagement with social robots and people's intention to use them (cf. H3, H4a, and H4b, see Table 2; Fig. 3). Because the exercises with Rob, Croq, and Car were heavily focused on physical tasks, we believe that evaluation of the robots was focused on whether the robot was an aid or obstacle in completing those tasks rather than their degree of realistic appearance. As a point of caution, this relationship is usually mediated by relevance and valence as shown in previous studies [39, 40]. In the current study, however, those factors were not taken into account. This absence of relevance and valence could be an explanation for why no significant direct effect of per-

Fig. 3 Tested hypotheses and their results, displayed as I-PEFiC relations



ceived affordances on use intentions was found (cf. H2, see Table 2; Fig. 3). Thus, the interaction effect that was found between perceived affordances and perceived aesthetics on use intentions should be considered with care.

Not each interrelation between the I-PEFiC factors (see Table 2; Fig. 3) was confirmed (direct effect of perceived affordances on use intentions, direct effect of perceived aesthetics on involvement and distance, and a direct effect of perceived realism on distance, see Table 2; Fig. 3). However, the results did indicate the existence of many of the predicted relationships. Perceived aesthetics (cf. H5a and H5b, see Table 2; Fig. 3) turned out to be not so influential for the engagement process, which is most likely due to the aesthetics being similar for each of the three social robots. LEGO enables the use of the same building blocks for different designs, which may result in a similar global aesthetics across the different designs. Nevertheless, the dominance of affordances, the moderate effects of aesthetics, and the modest role of realism are directly in line with I-PEFiC predictions.

Although (perceived) presence was not directly measured, participants frequently uttered phrases such as “Rob is the one for me” and “I liked doing the exercises together with Car” after completing the exercises. This may be an indication that the participants believed that they were interacting with another presence, which appears to be in line with a study by Kiesler et al., in which participants were aware that they were interacting with machines, yet still were engaged [19]. To validate whether the representation also has an effect on the perception of presence, future studies should consider also measuring social presence.

Yet, how form realism relates to the other dimensions of realism (behavioral realism, visual fidelity) still is difficult to define. In particular for designing robots, the current definition of realism and the accompanying (often contradictory) scientific evidence [46], makes it very difficult to determine how to appropriately design robot embodiments. Perhaps a different conceptualization of (form) realism is required. If form realism is the belief whether an embodiment

is believed as real or as something from fantasy, perhaps representation, as used by Bailenson et al. is a better term [2]. Although Bailenson et al. also appear to have confounded several dimensions of realism in their definition (“we have chosen to use the more general term representation type as a variable name and talk about the manipulation as mainly pertaining to photographic realism” [2, p. 33]), it may serve as the basis of our new conceptualization. If one considers realism from the perspective of representation, visual fidelity, and behavioral realism are subsets of representation. Each of these dimensions contributes to how much an embodiment refers to its corresponding real world entity (thus is experienced more or less realistic). This might be the reason why cartoons (low visual fidelity, high anthropomorphism, and high behavioral realism) still can be experienced as believable. Similarly, this also may account for why robots that are represented as realistic humans might be experienced as uncanny (high visual fidelity, high anthropomorphism, low behavioral realism, cf. the Uncanny Valley, [26]).

The focus should not be on a single degree of realism in a robot, but rather getting people into the experience as if interacting with something real and that the robot represents some part of reality (cf. [38]). However, this conceptualization is only a start, more research is required to define which parameters contribute to the belief that a robot is perceived as what it is designed to represent.

5 Limitations

Students from various locations were used; therefore, people of different age ranges or from different areas and educational backgrounds may respond differently to Rob, Croq, and Car. Furthermore, our study used convenience sampling to recruit participants, personal characteristics of the participants such as gender, prior robot experience, and age did not significantly affect the interaction or the engagement process.

The effects of the manipulated form realism might not have been found because the effect might be too small. If this is the case, 29 participants in a within-subjects design might not be enough to detect this effect. Future studies may want to improve the manipulation of form realism in an experimental design, by using other, more sophisticated materials to design the robots (whilst maintaining the same level of visual fidelity). Because the use of LEGO resulted in a similar look for each of the three robots, it could be that the differences between the social robots were too small. If the contrast between the robots is larger, designed form realism and perceived realism may result in a larger impact on participants' engagement and use intentions. Additionally, the lack of behavior might have resulted in similar evaluations of each robot.

Another issue that may have affected the results is that a talking humanoid robot is a better match in behavior and form compared to a talking vehicle or crocodile. During the debriefing some participants commented on their favorite robot. There was no clear favorite amongst the participants, some participants liked Croq, some liked Car, and others liked Rob. However, none of them indicated that a talking car or crocodile was weird or unnatural compared to a talking humanoid.

As each of the robots asked the participant to fill out the survey, there is a possibility that the robots have induced experimenter demand effects (cf. [29]). Because it was the robot asking the participant to evaluate it, participants might have displayed socially desirable behavior towards the robots. To prevent this, future studies should separate the interaction with the robot from the evaluation of the robot. Furthermore, this study only used subjective measures (self-report), but not behavioral measures. A meta-study demonstrated that subjective measures showed larger effect sizes compared to behavioral measures when measuring realism of agents and differences in behavior are not always detected by surveys [46]. For future studies that use I-PEFiC, it would be insightful to include both subjective and behavioral measurements to get to a more accurate prediction of engagement and intentions to use.

6 Conclusion

In all, the current findings revealed that designed form realism does not have an effect on perceived realism of differently shaped social robots when similar design features (LEGO Mindstorms) are used for different embodiments. For simple physical exercises like the ones used in this study, it is far more important that the robots work properly and help the user achieve his or her goal. This study provided further empirical evidence that the I-PEFiC model is an appropriate framework to apply in the domain of social robots to pre-

dict engagement and use intentions. The importance of our findings lies in the implication that in healthcare settings, designed form realism does not necessarily affect perceived realism, which may serve as an advice to social robots designers: A social robot embodiment that is designed to be more realistic (e.g., a nurse or a doctor) is not necessarily perceived as more realistic. It is what they afford, as well as their aesthetic appearance that makes them lovable and usable.

Acknowledgments This study is supported by the Creative Industry Scientific Programme (CRISP, Grant number: NWO 646.000.003). The research program is funded by a consortium of scientific and industrial partners, and by the Dutch Ministry of Education, Culture, and Science. We would like to thank the anonymous reviewers for the insightful feedback they provided, Marloes Spekman for her help and advice on statistics, Desmond Germans and Gideon Keyson for their advice related to using LEGO Mindstorms, and Wang Long Li and Allison Luger for their assistance during the experiment.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

1. Appel J, von der Pütten A, Krämer NC, Gratch J (2012) Does humanity matter? Analyzing the importance of social cues and perceived agency of a computer system for the emergence of social reactions during human-computer interaction. *Adv Human Comput Interact* 2012:13
2. Bailenson JN, Swinth K, Hoyt C, Persky S, Dimov A, Blascovich J (2005) The independent and interactive effects of embodied-agent appearance and behavior on self-report, cognitive, and behavioral markers of copresence in immersive virtual environments. *Presence* 14(4):379–393
3. Bailenson JN, Yee N, Merget D, Schroeder R (2006) The effect of behavioral realism and form realism of real-time avatar faces on verbal disclosure, nonverbal disclosure, emotion recognition, and copresence in dyadic interaction. *Presence* 15(4):359–372
4. Bartneck C, Kanda T, Mubin O, Al Mahmud A (2009) Does the design of a robot influence its animacy and perceived intelligence? *Int J Soc Robot* 1(2):195–204
5. Blascovich J, Loomis J, Beall AC, Swinth KR, Hoyt CL, Bailenson JN (2002) Immersive virtual environment technology as a methodological tool for social psychology. *Psycho Inq* 13(2):103–124
6. Broadbent E, Stafford R, MacDonald B (2009) Acceptance of healthcare robots for the older population: review and future directions. *Int J Soc Robot* 1(4):319–330
7. Broekens J, Heerink M, Rosendal H (2009) Assistive social robots in elderly care: a review. *Gerontechnology* 8(2):94–103
8. DiSalvo CF, Gemperle F, Forlizzi J, Kiesler S (2002) All robots are not created equal: the design and perception of humanoid robot heads. In: *Proceedings of the 4th conference on designing interactive systems: processes, practices, methods, and techniques*, pp. 321–326. ACM
9. Dobriarsky PJ, Suzman RM, Hodes RJ (2007) Why population aging matters—a global perspective. National Institute of Aging, Washington
10. Duffy BR (2003) Anthropomorphism and the social robot. *Robot Auton Syst* 42(3):177–190

11. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. *Robot Auton Syst* 42(3):143–166
12. Gibson JJ (1977) *The theory of affordances*. Lawrence Erlbaum, Hillsdale
13. Graham JW, Olchowski AE, Gilreath TD (2007) How many imputations are really needed? Some practical clarifications of multiple imputation theory. *Prev Sci* 8(3):206–213
14. Hayes AF (2013) *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. Guilford Press, New York
15. Hoorn JF (2012) *Epistemics of the virtual*. John Benjamins Publishing Company, Amsterdam
16. Hoorn JF, Konijn EA, Van der Veer GC (2003) Virtual reality: do not augment realism, augment relevance. *Upgrad Human Comput Interact* 4(1):18–26
17. Jung Y, Lee KM (2004) Effects of physical embodiment on social presence of social robots. In: *Proceedings of PRESENCE* pp. 80–87
18. Kidd CD, Breazeal C (2004) Effect of a robot on user perceptions. In: *Intelligent robots and systems, 2004. (IROS 2004)*. In: *Proceedings 2004 IEEE/RSJ international conference on*, vol. 4, pp. 3559–3564 IEEE
19. Kiesler S, Powers A, Fussell SR, Torrey C (2008) Anthropomorphic interactions with a robot and robot-like agent. *Soc Cognit* 26(2):169–181
20. Koda T, Maes P (1996) Agents with faces: the effect of personification. In: *Robot and human communication, 1996, 5th IEEE international workshop on*, pp. 189–194 IEEE
21. Konijn EA, Bushman BJ (2007) World leaders as movie characters? Perceptions of George W. Bush, Tony Blair, Osama Bin Laden, and Saddam Hussein. *Media Psychol* 9(1):157–177
22. Konijn EA, Hoorn JF (2005) Some like it bad: testing a model for perceiving and experiencing fictional characters. *Media Psychol* 7(2):107–144
23. Lee KM (2004) Presence, explicated. *Commun Theory* 14(1):27–50
24. Lee KM, Jung Y, Kim J, Kim SR (2006) Are physically embodied social agents better than disembodied social agents? The effects of physical embodiment, tactile interaction, and people's loneliness in human-robot interaction. *Int J Human Comput Stud* 64(10):962–973
25. MacDorman KF, Green RD, Ho CC, Koch CT (2009) Too real for comfort? Uncanny responses to computer generated faces. *Comput Human Behav* 25(3):695–710
26. Mori M (1970) The uncanny valley. *Energy* 7(4):33–35
27. Murray MK (2002) The nursing shortage: past, present, and future. *J Nurs Admin* 32(2):79–84
28. Nowak KL, Biocca F (2003) The effect of the agency and anthropomorphism on users' sense of telepresence, copresence, and social presence in virtual environments. *Presence* 12(5):481–494
29. Orne MT (1962) On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. *Am Psychol* 17(11):776
30. Von der Pütten AM, Krämer NC, Gratch J, Kang SH (2010) It doesn't matter what you are! explaining social effects of agents and avatars. *Comput Human Behav* 26(6):1641–1650
31. Reeves B, Nass C (1996) *How people treat computers, television, and new media like real people and places*. CSLI Publications and Cambridge University Press, Cambridge
32. Seyama J, Nagayama RS (2007) The uncanny valley: effect of realism on the impression of artificial human faces. *Presence* 16(4):337–351
33. Shapiro MA, Chock TM (2003) Psychological processes in perceiving reality. *Media Psychol* 5(2):163–198
34. Sidner CL, Lee C, Kidd CD, Lesh N, Rich C (2005) Explorations in engagement for humans and robots. *Artif Intell* 166(1):140–164
35. Sproull L, Subramani M, Kiesler S, Walker JH, Waters K (1996) When the interface is a face. *Human Comput Interact* 11(2):97–124
36. Super N (2002) Who will be there to care? The growing gap between caregiver supply and demand. *National Health Policy Forum*, Washington
37. Suzman RM, Beard J (2011) *Global health and aging*. National Institute of Aging, Bethesda
38. Van Vugt HC, Hoorn JF, Konijn EA (2009) Interactive engagement with embodied agents: an empirically validated framework. *Comput Anim Virtual Worlds* 20(2–3):195–204
39. Van Vugt HC, Hoorn JF, Konijn EA, de Bie Dimitriadou A (2006) Affective affordances: improving interface character engagement through interaction. *Int J Human Comput Stud* 64(9):874–888
40. Van Vugt HC, Konijn EA, Hoorn JF, Keur I, Eliëns A (2007) Realism is not all! User engagement with task-related interface characters. *Interact Comput* 19(2):267–280
41. Von Zitzewitz J, Boesch PM, Wolf P, Riener R (2013) Quantifying the human likeness of a humanoid robot. *Int J Soc Robot* 5(2):263–276
42. Wainer J, Feil-Seifer DJ, Shell DA, Mataric MJ (2006) The role of physical embodiment in human-robot interaction. In: *Robot and human interactive communication, 2006. ROMAN 2006. The 15th IEEE international symposium on*, pp. 117–122 IEEE
43. Wainer J, Feil-Seifer DJ, Shell DA, Mataric MJ (2007) Embodiment and human-robot interaction: a task-based perspective. In: *Robot and human interactive communication, 2007. RO-MAN 2007. The 16th IEEE international symposium on*, pp. 872–877 IEEE
44. Wang Y, Khooshabeh P, Gratch J (2013) Looking real and making mistakes. In: *Intelligent virtual agents*. Springer, Berlin Heidelberg
45. Wiese E, Wykowska A, Zwicker J, Müller HJ (2012) I see what you mean: how attentional selection is shaped by ascribing intentions to others. *PLoS one* 7(9):e45
46. Yee N, Bailenson JN, Rickertsen K (2007) A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp. 1–10 ACM

Robert A. Pauwe graduated with an engineering degree in interaction design (M.Sc. 2012, Delft University of Technology) and in industrial design engineering (B.Sc. 2008, Hogeschool Windesheim Zwolle). Currently, Robert is a Ph.D. candidate within the SELEMCA (Services of Electro-mechanical Care Agencies) project at VU University Amsterdam (NWO-funded CRISP-program). His Ph.D. research focuses on the design and realism of social robots and how these factors affect social interaction between robots and humans. His expertise is related to prototyping, interaction design, realism, and the education of designers. Additionally, he gives workshops in interaction design and supervises student projects at the VU University Amsterdam, Delft University of Technology, and the Glasgow School of Art.

Johan F. Hoorn graduated with two transdisciplinary PhD-theses: his first Ph.D. was in Literature and Psychology (VU University Amsterdam, 1997), while he obtained his second Ph.D. degree in Computer Science (VU University Amsterdam, 2006). He worked at Utrecht University, Tilburg University, and in four different schools at VU University Amsterdam (Humanities, Science, Life Sciences, and Social Science). Johan is an adjunct professor of the Hong Kong Polytechnic University and the former director of the Center for Advanced Media Research Amsterdam at VU University. Currently, Johan is the principal investigator and project leader of SELEMCA, a multi-million, multi-

stakeholder research and design project in social robotics for the care domain, granted by the Ministry of Education, Culture, and Science.

Elly A. Konijn (Ph.D. 1994, Utrecht University i.c.w. University of Amsterdam) is a Full Professor of Media Psychology at the Department of Communication Science and a Fenna Diemer-Lindeboom endowed chair at VU University Amsterdam. She has graduated in Psychology, Social Scientific Information Systems, and Media Studies, and was a visiting scholar at the City University of New York. Her current scientific research program moves along three main lines: 1) media use among adolescents (e.g., cyberbullying, violent video game effects, thin-body ideal); 2) emotions and reality-fiction perceptions in entertainment media (e.g., how we learn from fiction, morality issues); and 3) virtual humans in virtual environments (e.g., new technology applications, affective bonding with robots). She is past chair of the Information Systems division and board member of the International Communication Association and former editor of the ISI-journal *Media Psychology*.

David V. Keyson is a Full Professor and heads the research program in Sustainable Living and Work at the Faculty of Industrial Design Engineering at Delft University of Technology. His educational work focuses on interactive technology design in the context of smart products and environments. Prior to joining the Delft University of Technology, he worked at Philips Research as a Senior Research Scientist in Media Interaction and at Xerox in the department of Industrial Design and Human Interface in California. He holds a Ph.D. (1996) from the Technical University of Eindhoven in Perception and Technology and a degree in Ergonomics (M.Sc. 1987) from Loughborough University.