

## Attitude Towards Humanoid Robots and the Uncanny Valley Hypothesis

Paweł Lupkowski, Marta Gierszewska \*

**Abstract.** The main aim of the presented study was to check whether the well-established measures concerning the attitude towards humanoid robots are good predictors for the uncanny valley effect. We present a study in which 12 computer rendered humanoid models were presented to our subjects. Their declared comfort level was cross-referenced with the Belief in Human Nature Uniqueness (BHNU) and the Negative Attitudes toward Robots that Display Human Traits (NARHT) scales. Subsequently, there was no evidence of a statistical significance between these scales and the existence of the uncanny valley phenomenon. However, correlations between expected stress level while human-robot interaction and both BHNU, as well as NARHT scales, were found. The study covered also the evaluation of the perceived robots' characteristic and the emotional response to them.

**Keywords:** Uncanny Valley Hypothesis, human-likeness, computer-generated models, attitude towards robots, Belief in Human Nature Uniqueness (BHNU), Negative Attitudes toward Robots that Display Human Traits (NARHT), HRI, social robotics

### 1. Introduction

The Uncanny Valley Hypothesis (hereafter UVH) has been formulated by Masahiro Mori (see [13]). Mori hypothesises that when we present a subject with a series of different human-like models (including robots) certain models will trigger negative reactions (uneasiness, eeriness). As he claims these will be almost human-like characters. We may imagine models presented in order on the  $X$  axis—from the least human-like (like e.g., robotic arm) to the most human-like ones. On the  $Y$  axis we would present the affinity level. According to Mori's suggestion we would observe

---

\*Department of Logic and Cognitive Science, Institute of Psychology, Adam Mickiewicz University, Poznań; Reasoning Research Group {pawel.lupkowski, marta.gierszewska}@gmail.com

the growing level of affinity as we move towards human-like models, but on a certain degree of human-likeness the level of affinity will rapidly get lower, this is the ‘valley’ (or as MacDorman and Norri Kageki call it ‘descent into eeriness’ [13]).

As such UVH is clearly applicable in fields of social robotics and human-robot interaction (HRI)—see e.g.[1]. However, as many studies show, UVH also plays its role for game design and computer animation (see e.g. [21], [10] and [4]). Research presented in [5] brings empirical evidence for the existence of the uncanny valley in animated film characters (which was to this point suggested in popular media, e.g. for the ‘Polar Express’ or ‘Beowulf’ animations<sup>1</sup>). Our study addresses this field of UVH applicability. We have presented our subjects with 12 computer rendered models ranging from simplistic robots, through cartoonish characters to realistic humans. The research aim was to check whether the well-established measures concerning the attitude towards humanoid robots may serve as predictors for the uncanny valley (hereafter UV) effect. The main motivation for such a research question is the need of deep analysis of the UVH pointed out in the exhaustive review of the hypothesis presented in [6]. The authors of the aforementioned paper point out that empirical results concerning the uncanny valley are ambiguous. Among many reasons for this, one seems to be especially important—namely the operationalisation of the term used to describe *Y* axis of the UVH graph. Authors use different concepts to grasp the idea behind this dimension of interacting with humanoid robots—like, e.g. ‘comfort level’ in [20], ‘likeability’ in [11] or ‘acceptability’ in [16]. As we read in [6]: “Given that the original terms for the affinity dimension (or at least their common translations) are ambiguous, empirical studies would be necessary for resolving which self-report items would be ideal for measuring affinity.”. Staying on the level of subjects’ declarations concerning presented stimuli we aimed at well-established questionnaires concerning human attitude towards humanoid robots. We have chosen two such questionnaires presented in [17]. These were: the Belief in Human Nature Uniqueness (BHNU) and the Negative Attitudes toward Robots that Display Human Traits (NARHT)—both in Polish adaptations. Our expectation was that results in these questionnaires may serve as a predictor for the uncanny valley effect, so that these scales may be useful in operationalisation of *Y* axis dimension and shed some light on the UVH studies’ results.

The structure of the paper is following. In Section 2 we describe in detail the set of stimuli used for the study and the tools we use, namely NARHT and BHNU questionnaires. We also present the research procedure and hypotheses. Section 3 presents results of the study. The last section covers summary and discussion of results. The paper is supplemented with an Appendix, where all models used for the study are presented.

---

<sup>1</sup>Paul Clinton in his review of ‘The Polar Express’ [2] writes: “The overall artwork is remarkable, and the action sequences are inventive and emotionally gripping. [...] But those human characters in the film come across as downright... well, creepy.” David Gallagher wrote in the similar manner about ‘Beowulf’ [3]: “The movie definitely pushes digital acting far beyond anything I’ve seen before—but it looks as if the last few yards of the journey toward convincing realism are going to be the really hard part. [...] When it was over, I felt relieved to be back in the company of uncreepy flesh-and-blood humans again.”

## 2. Methods

### 2.1. Stimuli

For our research we have used 12 computer-generated models. Models were retrieved from 3D character banks<sup>2</sup> and rendered with the use of the *Unity* environment<sup>3</sup>. All models were chosen arbitrary by the authors and then consulted with two designers experienced in the game development. Models covered various levels of degrees of human-likeness (DOH)—ranging from simplistic robots, through androids, animated characters to human models. Key features which we took into account were: visible facial features, detail level of the model (e.g. visible hands and fingers), overall style of the model. For example, models intended as simplified robots do not have visible eyes and their hands and feet are not detailed. Also, joints of body parts are clearly visible, which give them a very mechanical appearance. These features were previously tested in a study concerning degrees of human likeness presented in [7]. On the basis of this previous study we have resigned from zombie models in the stimuli set as they triggered strong emotional reactions. Before our research, we have consulted the stimuli set used for the study with respect to this issue with the Reasoning Research Group members (during two seminar meetings). All models used in the study are presented in the Appendix of this paper. Models have the same height and are presented front face in a neutral pose on a neutral background (empty room). This allows evaluation of body proportions, body elements (such as hands, feet) and even facial expressions.

### 2.2. Questionnaires

For the study we have decided to use a sub-scale of the Negative Attitudes Towards Robots Scale (NARS). The scale was first described in [14] and its Polish adaptation (NARS-PL) is presented in [17]. NARS measures psychological reactions to human-like and non-human-like robots. The focus is put on the extent to which one would be reluctant to interact with a robot—see [17, p. 65]. The reliability results for the adaptation study of the analysed scales (on 213 subjects) are satisfactory and as the authors of the study write: “It is argued that NARS-PL is a useful tool to predict human responses to social robots in HRI studies in Poland.” [17, p. 70]. As such, this tool fits well into our research goals.

We have not used the full questionnaire, but only the following items form the NARS sub-scale called the Negative Attitudes toward Robots with Human Traits (NARHT):

- *I would feel uneasy if robots really had emotions.*
- *Something bad might happen if robots developed into living beings.*

<sup>2</sup>See <http://tf3dm.com> and <https://www.mixamo.com/>.

<sup>3</sup><https://unity3d.com/>.

- *If robots had emotions, I would be able to make friends with them.*
- *I feel comforted being with robots that have emotions.*
- *I would hate the idea that robots or artificial intelligences were making judgments about things.*

The reason for using the shortened version of the tool was that we wanted our research to last for a possibly short time—as it was conducted *online*. Participants responded on a 5-point scale (1 – totally disagree to 5 – totally agree).

The aim of NARHT is to capture the responses to robots that display human traits like emotions, language, and agency.

We have also added one question addressing directly the issue of human-robot interaction, namely:

- *I would feel very nervous just standing in front of a robot.*

For the study we also used the Belief in Human Nature Uniqueness scale (BHNU). BHNU aims at assessing the extent to which humans reserve human nature for their own group and deny the possibility of a human essence to robots (see [17, p. 67]). BHNU consists of the following questions (we have resigned from one BHNU question in our research, namely *Even if ultra-sophisticated a robot will never be considered as human being.*):

- *Even if ultra-sophisticated...*
  1. *a robot will never feel the same emotions as a human being,*
  2. *a robot will never use language in the same way as a human being;*
  3. *a robot will always be a mechanical imitation of the human being;*
  4. *a robot will never have consciousness;*
  5. *a robot will never have morality.*

Also for this case, participants responded on a 5-point scale (1 – totally disagree to 5 – totally agree).

NARHT and BHNU scales were supplemented with additional questions displayed along the computer rendered models as it is described in the following subsection.

### 2.3. Procedure and subjects

The research was conducted as an *on-line* questionnaire with the use of Google Forms tool. The data was collected during the November 2017. The language of the study was Polish. Subjects were recruited *via* email invitations and announcements published on popular social media web-pages. In the study 100 participants took part (55 women and 44 men; 1 subject refused to provide this information), mean age of a participant was 27.78 ( $SD = 8.28$ ).

First, a subject was presented with the instruction, explaining the procedure. Subjects were informed about anonymity and the possibility of withdrawing from the research at any time. After reading this information, the subject pressed a key and started the questionnaire.

In the first part of the study, subjects filled out a questionnaire consisting of NARHT and BHNU questions. The second part, consisted of presenting 12 models. Each model was presented separately. Under the picture the following questions were displayed (below we present their translation, as the study was in Polish):

1. *How much does the presented model resemble a human?* (Answers on a scale 1–5, where: (1) Completely not human-like; (2) Rather not human-like; (3) It starts to look human-like; (4) Rather human-like; (5) Completely human-like).
2. *Mark on a given scale to which extent you would agree that the presented model is intelligent.* (Answers on a 5-point scale: 1 – totally disagree to 5 – totally agree).
3. *Mark on a given scale to which extent you would agree that the presented model is trustworthy.* (Answers on a 5-point scale: 1 – totally disagree to 5 – totally agree).
4. *Mark on a given scale to which extent you would agree that the presented model is hostile.* (Answers on a 5-point scale: 1 – totally disagree to 5 – totally agree).
5. *Mark on a given scale to which extent you would agree that the presented model is strong.* (Answers on a 5-point scale: 1 – totally disagree to 5 – totally agree).
6. *How comfortable are you when you watch this model in a given environment?* (Answers on a scale 1–5, where (1) very comfortable, (2) quite comfortable, (3) neutral, (4) uncomfortable, (5) very uncomfortable).
7. *What is your emotional reaction for the presented model?* (Four possible answers: (i) the model seems to be friendly; (ii) my reaction is neutral; (iii) the model looks strange; (iv) the model makes me feel anxious. Subjects could choose maximally two answers.).

At the end of the study, subjects were asked to answer two more questions: about their age and gender.

## 2.4. Hypotheses

The main aim of the study was to verify whether we can predict appearance of the uncanny valley effect on the basis of NARHT and/or BHNU results. Thus, our hypotheses were the following.

- (H1) The UV effect appearance will correlate with the NARHT result.
- (H2) The UV effect appearance will correlate with the BHNU result.

- (H2a) Subjects will experience the UV effect for models for which they will ascribe a combination of the following human features: intelligence, strength, hostility or the lack of trustworthiness.

### 3. Results

For the data analysis we used R statistical software [18] (version 3.3.1) and PS IMAGO 4.0 (IBM SPSS Statistics).

The (slightly modified) NARHT and BHNU scales used in our study had a satisfactory psychometric features. The reliability (as expressed by the Cronbach's alpha) is  $\alpha = 0.77$  for the NARHT and  $\alpha = 0.84$  for the BHNU.

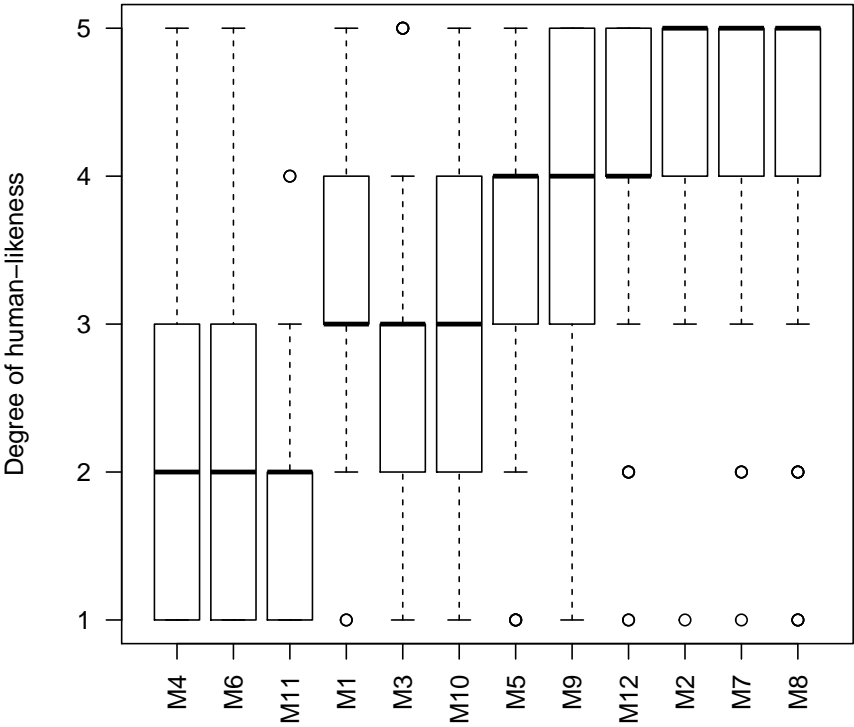
Before checking research hypotheses we have to analyse for which models the UV effect was observed. First, we order models accordingly to their degrees of human-likeness (as assessed by subjects). After that we can analyse the declared interaction comfort and emotional reaction to models from our study set.

**Degrees of human-likeness** On the basis of subjects' evaluation we can group 12 models into four groups with growing human-likeness levels—see Figure 1. The first group (models M4, M6 and M11—numbers of models indicate the order in which they appeared in the study) were evaluated as 2 (median), i.e. 'rather not human-like'. Next group (M1, M3, M10) was evaluated as 'it starts to look human-like' (median=3). Next to the right we have cartoon-like models (M5, M9, M12) evaluated as 'rather human-like' (median=4). The rightmost group consists of 'completely human-like' (median=5) models M2, M7 and M8. It is worth to notice that none of the models received the median score lower than 2 (which would mean that the model was evaluated as 'completely not human-like').

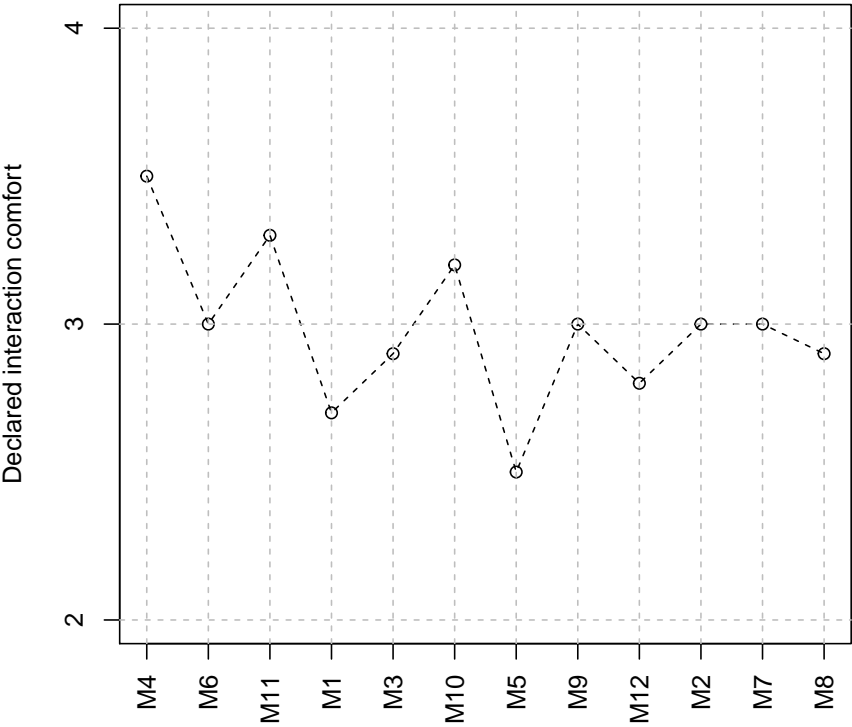
**Uncanny valley** After ordering the models accordingly to their DOH we may ask a question, whether the uncanny valley may be observed in our study. The declared comfort level for the models is presented in Figure 2. In the questionnaire answer (1) meant *very comfortable* and (5) *very uncomfortable*, consequently the uncanny valley is visible in the figure as a raise not decrease on the Y axis.

The highest comfort level was declared for M5 (cartoon-like character). The lowest comfort level was declared for M4 (robot). We observe a small UV effect (the lower declared comfort level for a given model compared to models on its left and right side) for four models. These are: M11, M3 and M10 and M9. These models are presented in Figure 3.

**Emotional reaction** In the study we have also asked about a declared emotional reaction for models. Subjects answered the question: *What is your emotional reaction for the presented model?*. Among four possible answers ((i) the model seems to be friendly; (ii) my reaction is neutral; (iii) the model looks strange; (iv) the model makes



**Figure 1.** Human-likeness assessment of models used in the study. The ordering is accordingly to growing human-likeness. Numbers of models indicate the order in which they appeared in the study

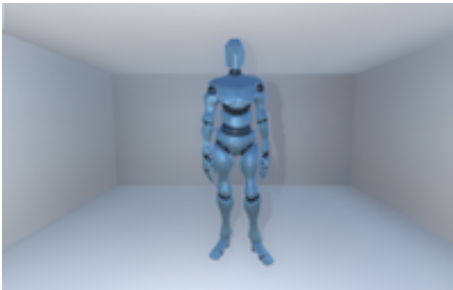


**Figure 2.** Mean declared interaction comfort for models. The ordering of the models is established on the basis of subject’s assessment. Scale for comfort is reversed, so a ‘valley’ is actually visible as a peak in the figure

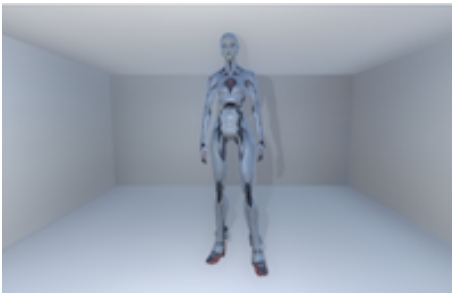




M11



M3



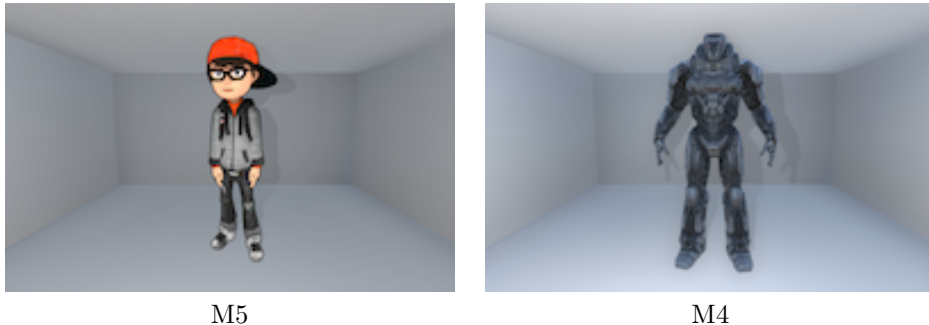
M10



M9

**Figure 3.** Models for which a (small) UV effect was observed for the declared interaction comfort

me feel anxious) subjects could choose maximally two. For M5 (see Figure 4) 63% of subjects decided that the model seems to be friendly. It is also the model with the highest level of declared interaction comfort. We observe analogous correspondence for M4 (battle robot, see Figure 4). For this model we observe the lowest declared interaction comfort level and 76% of subjects declared that the model makes them feel anxious.



**Figure 4.** Models M5 and M4. M5 received the highest declared interaction comfort score and was assessed as the ‘friendly’ one by subjects. M4 received the lowest declared interaction comfort score and was assessed as the ‘anxious’ one

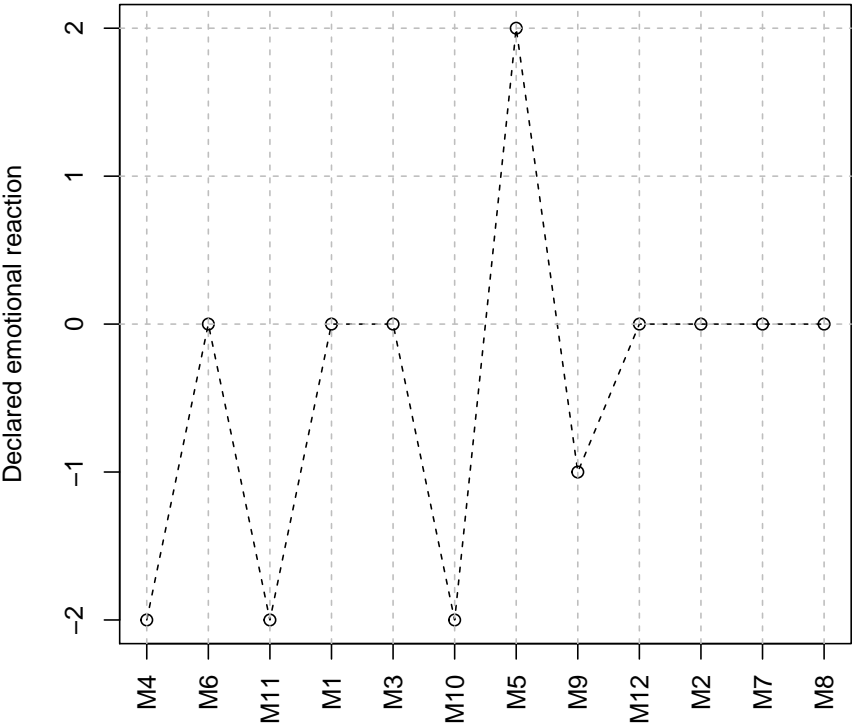
Let us now take a closer look on the emotional reaction to models with the observed UV effect. Starting from the leftmost model M11, the dominant reaction was anxiety (44%). 13% of subjects declared a combination of two reactions: ‘the model looks strange’ and ‘the model makes me feel anxious’. For the model on the right to M11, namely M3 more than a half of subjects (57%) declared that their reaction is neutral. For the next model, i.e. M10 subjects declared anxiety (39%) strangeness of the model (38%) or both combined (7%). For model M9 39% of subjects declared that it looks strange. This may be the effect of certain imperfections (halo effect) of the rendered model. The summary of dominant declarative emotional reactions is presented in Table 1.

Graphical representation of the emotional reaction for our models is presented in Figure 5. For this graph we encoded reactions accordingly to the following: neutral = 0, strange = -1, anxiety = -2, friendly = 2.

After the analysis of the UV effect and emotional reaction for the set of our models we may evaluate research hypotheses.

### **Hypothesis H1: The UV appearance will correlate with the NARHT result**

For establishing the attitude towards humanoid robots score the mean for answers to 6 questions of the NARHT scale were taken into account (see [17]). The r-Pearson test was used in order to check the correlation between the declared comfort level and NARHT score. No significant correlation was observed for all 12 models ( $\rho = 0.08$ ;  $p = 0.44$ ). When we check the correlation of the declared comfort level and the



**Figure 5.** Declared emotional reaction for models. The ordering of models is established on the basis of subject’s assessment. -2 = anxiety, -1 = strange, 0 = neutral, 2 = friendly.

**Table 1.** The summary of dominant declarative emotional reactions for models

Model	The dominant emotional reaction	% of declarations
M1	neutral	65%
M2	neutral	45%
M3	neutral	57%
M4	anxiety	76%
M5	friendly	63%
M6	neutral	39%
M7	neutral	55%
M8	neutral	50%
M9	strange	39%
M10	anxiety	39%
M11	anxiety	44%
M12	neutral	51%

NARHT score for separate models we observe significant results for M1 ( $\rho = 0.267$ ;  $p = 0.007$ ), M4 ( $\rho = 0.327$ ;  $p = 0.001$ ) and M6 ( $\rho = 0.287$ ;  $p = 0.004$ ). We also observe a negative correlation for model M5 ( $\rho = -0.265$ ;  $p = 0.008$ ). None of these models triggered UV effect in our research—consequently, we cannot confirm hypothesis H1.

**Hypothesis H2: The UV appearance will correlate with the BHNU result**  
The BHNU score was established as a mean of answers to 5 questions to BHNU questionnaire (see [17]). For the analysis of correlation between the BHNU score and the declared comfort also the r-Pearson test was used. No significant correlation was observed for all 12 models ( $\rho = 0.078$ ;  $p = 0.44$ ). The detailed analysis of each model revealed a significant negative correlation for model M5 ( $\rho = -0.223$ ;  $p = 0.026$ ). This means that for this model the higher belief of the uniqueness of the human nature, the more comfortable a subject felt during M5 exposition. As a result we cannot confirm hypothesis H2.

At this point we also receive an answer to our main research question. As both H1 and H2 are not confirmed we cannot simply use NARHT and BHNU scales to predict the UV effect appearance.

**Hypothesis H2a: Subjects will experience the UV effect for models for which they will ascribe a combination of the following human features: intelligence, strength, hostility or the lack of trustworthiness** For H2a evaluation we used medians from the answers to questions concerning models’ features. Detailed results are presented in Table 2. Let us now take a closer look on the features ascribed to models for which UV effect has been observed (M3, M9, M10, M11).

A distinctive feature of M3 is *strength* (median=4). Other features were not relevant to the model accordingly to our subjects (medians on the level of 3 may be interpreted as such evaluation). As for M9, what is interesting is that it received low

**Table 2.** A comparison of models’ features

Model	Intelligent	Trustworthy	Hostile	Strong	UV
M1	3.0	3.0	2.0	3.0	no
M2	3.0	3.0	2.0	3.0	no
M3	3.0	3.0	3.0	4.0	yes
M4	3.0	2.0	4.0	5.0	no
M5	3.0	3.0	1.5	2.0	no
M6	3.0	3.0	3.0	3.0	no
M7	3.0	3.0	2.0	3.0	no
M8	3.0	3.0	3.0	3.0	no
M9	3.0	3.0	2.0	2.0	yes
M10	4.0	3.0	3.0	4.0	yes
M11	2.0	2.0	3.0	4.0	yes
M12	3.0	3.0	3.0	3.0	no

marks (median=2) for features as *hostility* and *strength*. M10 received the highest scores for the *intelligence* feature in the whole study sample. The last model with observed UV effect, M11 received high score (median=4) for *strength* and low (median=2) for *intelligence*. On the basis of these assessments we cannot confirm H2a as there is no clear pattern of ascribed features in the case of models with observed UV effect.

Additionally, we have also performed a correlation analysis for the models’ features and declared interaction comfort. The details are presented in Table 3.

**Table 3.** Correlation of models’ features and declared interaction comfort (\* means that a correlation is significant)

Declared interaction comfort and								
Model	intelligence		trustworthiness		hostility		strength	
	<i>rho</i>	<i>p</i>	<i>rho</i>	<i>p</i>	<i>rho</i>	<i>p</i>	<i>rho</i>	<i>p</i>
M1	.139	.167	-.131	.195	.162	.108	.006	.956
M2	-.186	.064	-.340*	.001	.200*	.046	-.017	.864
M3	.015	.884	-.135	.181	.338*	.001	.225*	.024
M4	-.027	.793	-.329*	.001	.458*	.000	.394*	.000
M5	-.218*	.029	-.223	.026	.236*	.018	-.075	.456
M6	.085	.401	-.266*	.007	.469*	.000	.174	.084
M7	-.170	.092	-.322*	.001	.196	.050	-.067	.507
M8	-.370*	.000	-.521*	.000	.345*	.000	.024	.809
M9	-.373*	.000	-.379*	.000	.270*	.007	.021	.837
M10	.062	.107	-.437*	.000	.402*	.000	.148	.143
M11	-.096	.343	-.339*	.001	.366*	.000	.133	.187
M12	-.166	.100	-.417*	.000	.177	.078	-.070	.489

As it may be observed there is a strong correlation between *trustworthiness* of a model and declared interaction comfort (the more trustworthy the model is, the higher level of comfort is declared). Such a correlation is observed for 75% of models from the study sample. Analogically, for 75% of models correlation between *hostility* and

comfort is observed (the more hostile the model, the lower level of comfort is declared). As for the *intelligence* feature and comfort correlation is observed for three models. What is somehow surprising, when it comes to *strength* and comfort, we observe a significant correlation only for two models.

Let us now analyse the last question asked in our survey—namely the one asking about a direct interaction with a robot: *I would feel very nervous just standing in front of a robot* (which was presented together with NARHT scale). Answers to this question correlate ( $\rho = 0.27$ ;  $p = 0.007$ ) with the declared comfort level. It is worth to mention that answers to this question also correlate ( $\rho = 0.366$ ,  $p < 0.00$ ) with a BHNU scale. The higher the belief in human uniqueness is, the more nervous subjects are when thinking about meeting a robot. This result is in line with the one reported in [17].

#### 4. Summary and discussion

The aim of the described study was to validate whether we can predict the UV effect using certain well-established scales. These were the Belief in Human Nature Uniqueness (BHNU) and the Negative Attitudes toward Robots that Display Human Traits (NARHT) in their Polish adaptations presented in [17].

Accordingly to H1, we should observe a correlation of NARHT results and declared interaction comfort. Such a correlation was not observed for models for which the UV effect appeared. As for H2, our results do not confirm that there exist a significant correlation between BHNU scale and declared comfort levels for models with the UV effect observed. This brings us to the conclusion that NARHT and BHNU cannot be simply used as the UV effect predictors. However, we should point out one important issue with the reported research. Namely, the UV effect for our study sample was observed for four models but it was not very strong (see Figure 2). This indicates a need for further studies with the use of NARHT and BHNU scales in the context of the UVH.

Possibly also modifications and improvements of models from the study sample are required. In our opinion analysis of H2a constitute a good starting point for such process. No clear pattern of ascribed features for models with the UV effect was observed, but we obtained a valuable results about relations of individual features and a declared interaction comfort. Features as trustworthiness and hostility are especially important in the light of our findings. It is worth to mention here that when it comes to human-likeness assessment of models from our study sample we have obtained analogical results to the ones from the study presented in [7]. What is more, the same set of models was used in another study concerning the background condition for UVH (see [8])—also in this case the DOH ordering was the same (four groups of models were established by subjects' evaluations). This suggests that the set of models prepared for aforementioned studies provides a good and already tested starting point for future modifications and research (see the internal validity discussion for UVH related research in [6]).

A positive result was observed for the additional question directly addressing a stress related to meeting a robot. For this question correlation with declared comfort is observed. This result is promising for designing future research related to UVH. The observed correlation is in line with studies reporting that individual attitudes toward robots influence how subjects interact with robots [19, 15, 17, 9]. The correlation between the aforementioned question and the BNHU result is also in line with results of previous studies [17, 9]. The higher a belief in human uniqueness is, the more nervous subjects are when thinking about meeting a robot. This result may be interpreted in the light of categorical approach to the uncanny valley hypothesis [6, 22, 12]. Subject who belief that humans are unique probably have a clear categorisation of human and non-human entities. When confronted with a model which does not belong clearly to one of these categories, discomfort appears.

## Acknowledgment

The authors give their thanks to Dagmara Dziedzic and Wojciech Włodarczyk for their help in preparing models for the study. We would also like to give our thanks to Marcin Jukiewicz for helpful feedback and comments on a draft of this paper.

## References

- [1] Becker-Asano C., Ogawa K., Nishio S., and Ishiguro H. Exploring the uncanny valley with geminoid hi-1 in a real-world application. In *Proceedings of IADIS International conference interfaces and human computer interaction*, pages 121–128, 2010.
- [2] Clinton P. Review: ‘Polar Express’ a creepy ride. technology brilliant, but where’s the heart and soul?, November 2004. <http://edition.cnn.com/2004/SHOWBIZ/Movies/11/10/review.polar.express/>.
- [3] Gallagher D. F. Digital actors in ‘Beowulf’ are just uncanny, November 2007. <https://bits.blogs.nytimes.com/2007/11/14/digital-actors-in-beowulf-are-just-uncanny/>.
- [4] Geller T. Overcoming the uncanny valley. *IEEE computer graphics and applications*, 28(4), 2008.
- [5] Kätsyri J., Mäkräinen M., and Takala T. Testing the ‘uncanny valley’ hypothesis in semirealistic computer-animated film characters: An empirical evaluation of natural film stimuli. *International Journal of Human-Computer Studies*, 97:149–161, 2017.
- [6] Kätsyri J., Förger K., Mäkräinen M., and Takala T. A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Frontiers in Psychology*, 6:390, 2015.

- [7] Łupkowski P., Rybka M., Dziedzic D., and Włodarczyk W. Human-likeness assessment for the uncanny valley hypothesis. *Bio-Algorithms and Med-Systems*, 13(3):125–131, 2017.
- [8] Łupkowski P., Rybka M., Dziedzic D., and Włodarczyk W. The background context condition for the uncanny valley hypothesis. *International Journal of Social Robotics*, Sep 2018. Online first. 10.1007/s12369-018-0490-7.
- [9] MacDorman K. F. and Entezari S. O. Individual differences predict sensitivity to the uncanny valley. *Interaction Studies*, 16(2):141–172, 2015.
- [10] MacDorman K. F., Green R. D., Ho C.-C., and Koch C. T. Too real for comfort? uncanny responses to computer generated faces. *Computers in human behavior*, 25(3):695–710, 2009.
- [11] Maya R. Navigating a social world with robot partners: A quantitative cartography of the uncanny valley. *146*, 2:22–32, 2016.
- [12] Moore R. K. A bayesian explanation of the ‘uncanny valley’ effect and related psychological phenomena. *Scientific reports*, 2:864, 2012.
- [13] Mori M., MacDorman K. F., and Kageki N. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine*, 19(2):98–100, 2012. (Original work published in 1970 in Japanese).
- [14] Nomura T., Kanda T., and Suzuki T. Experimental investigation into influence of negative attitudes toward robots on human–robot interaction. *Ai & Society*, 20(2):138–150, 2006.
- [15] Nomura T., Kanda T., Suzuki T., and Kato K. Prediction of human behavior in human–robot interaction using psychological scales for anxiety and negative attitudes toward robots. *IEEE transactions on robotics*, 24(2):442–451, 2008.
- [16] Piwek L., McKay L. S., and Pollick F. E. Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, 130(3):271–277, 2014.
- [17] Pochwatko G., Giger J.-C., Róžańska-Walczyk M., Świdrak J., Kukielka K., Możaryn J., and Piçarra N. Polish version of the negative attitude toward robots scale (NARS-PL). *Journal of Automation Mobile Robotics and Intelligent Systems*, 9(3):65–72, 2015.
- [18] R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria, 2013. acess 20.03.2017.
- [19] Rosenthal-von der Pütten A. M. and Krämer N. C. How design characteristics of robots determine evaluation and uncanny valley related responses. *Computers in Human Behavior*, 36:422–439, 2014.

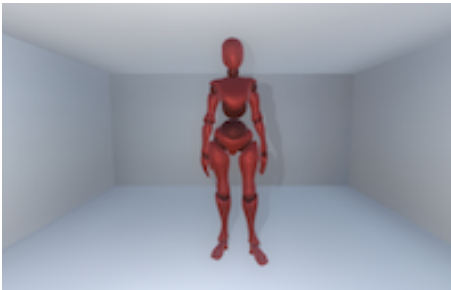


- [20] Seyama J. and Nagayama R. S. The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoper. Virtual Environ.*, 16(4):337–351, Aug. 2007.
- [21] Ueyama Y. A bayesian model of the uncanny valley effect for explaining the effects of therapeutic robots in autism spectrum disorder. *PloS one*, 10(9):e0138642, 2015.
- [22] Yamada Y., Kawabe T., and Ihaya K. Categorization difficulty is associated with negative evaluation in the “uncanny valley” phenomenon. *Japanese Psychological Research*, 55(1):20–32, 2013.

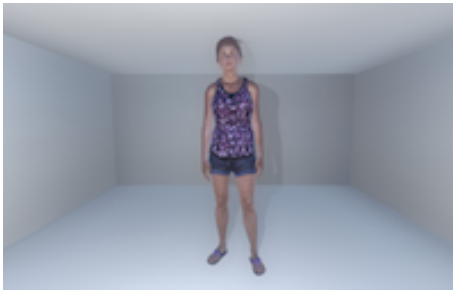
*Received 13.08.2018, Accepted 21.11.2018*

Appendix

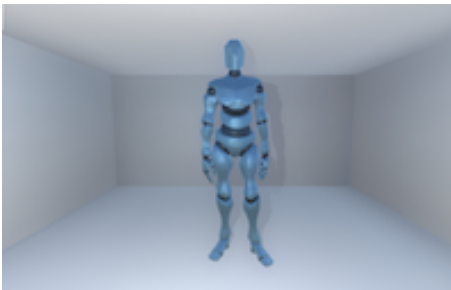
Numbers of models indicate the order in which they appeared in the study.



M1



M2



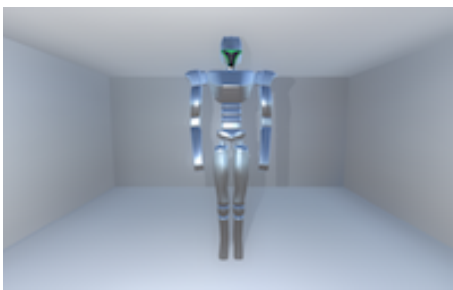
M3



M4



M5



M6



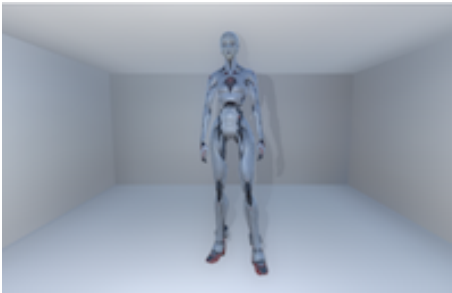
M7



M8



M9



M10



M11



M12