



The Background Context Condition for the Uncanny Valley Hypothesis

Paweł Łupkowski¹ · Marek Rybka² · Dagmara Dziejczak³ · Wojciech Włodarczyk⁴

Accepted: 26 August 2018 / Published online: 8 September 2018
© The Author(s) 2018

Abstract

We present the results of the background context condition experiment for the uncanny valley hypothesis. Subjects were presented with 12 computer-generated 3D models in two background variants. For the first group 12 models were rendered on a neutral background (empty room) and for the second group the same models were rendered on a suitable background, which was relevant to a given model (science-fiction scenery, town etc.). The aim of this study was to check whether the background context would influence differences in comfort level, human-likeness rating and emotional reaction to the models. A statistically significant difference in comfort levels was observed only for one of the models, the same situation was noticed for emotional reactions. We also tested the possibility of existence of more than one uncanny valley related to different factors as suggested in Kätsyri et al. (Front Psychol 6:390, 2015. <https://doi.org/10.3389/fpsyg.2015.00390>).

Keywords Uncanny valley hypothesis · Human-likeness · Categorisation · Computer-generated models · Background context

1 Introduction

The uncanny valley hypothesis (hereafter UVH) has been formulated by Mori [9]. 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, from the least human-like (like e.g., robotic arm) to the most human-like ones on the X axis. On the Y axis we would present the affinity level. According to Mori's suggestion we would observe a growing level of affinity as we move towards human-like models, but on a certain level 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'—[9]).

As such, UVH gained much attention in both, research and popular media. Kätsyri et al. [5] note that the reason for this is the wide range of applicability and importance of this hypothesis for a wide range of disciplines, from social robotics to computer animation and game design (see e.g. [12] and an overview in [2,7]). As the widely discussed case of the review of 'Polar Express' shows, UVH may have an effect not only when we interact with real robots, but also for a computer animation domain (see the evidence for the existence of uncanny valley in animated film characters presented in [4]). Paul Clinton 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. So 'The Polar Express' is at best disconcerting, and at worst, a wee bit horrifying."¹

Kätsyri et al. [5] in their exhaustive review of different theoretical models of UVH and their up-to date empirical support note that "it is surprising that empirical evidence for the uncanny valley hypothesis is still ambiguous if not non-existent". Their main conclusion is that the need of

✉ Paweł Łupkowski
Pawel.Lupkowski@amu.edu.pl

¹ Department of Logic and Cognitive Science, Institute of Psychology, Reasoning Research Group, Adam Mickiewicz University, ul. H. Wieniawskiego 1, Poznan, Poland

² Institute of Psychology, Reasoning Research Group, Adam Mickiewicz University, ul. H. Wieniawskiego 1, Poznan, Poland

³ Department of Logic and Cognitive Science, Institute of Psychology, Adam Mickiewicz University, ul. H. Wieniawskiego 1, Poznan, Poland

⁴ Faculty of Mathematics and Computer Science, Adam Mickiewicz University, ul. H. Wieniawskiego 1, Poznan, Poland

¹ 2004 'Review: 'Polar Express' a creepy ride. Technology brilliant, but where's the heart and soul?' by Paul Clinton for CNN.com (<http://edition.cnn.com/2004/SHOWBIZ/Movies/11/10/review.polar.express/>, Access 4.12.2017).

deep analysis of the hypothesis itself, which would include inspection of human-likeness and affinity dimensions of the UVH. Especially as Kätsyri et al. [5] notice, Mori “himself used anecdotal examples to characterise different degrees of human-likeness” using industrial robots, toy robots, prosthetic and even puppets, corpses and zombies. Such a variety of proposed models may trigger various levels of evaluating human-likeness when it comes to empirical studies. In line with these observations we have designed a study which involves tracking not only comfort of interaction with a given model but also the degree of human-likeness (DOH), difficulty of assessing DOH, and subject’s emotional reaction when confronted with the model.

However, our main research question for this study was whether a background context change will play a role in the aforementioned dimensions. For this we have used 12, previously tested (see [6]) computer-generated models. We have prepared two sets of stimuli: (A) models presented on a neutral background (empty room) and (B) the same models rendered on a suitable background, which was relevant to a given model (science-fiction scenery, town etc.). The inspiration for this study was threefold. Firstly it came Hanson’s [3] critic of the uncanny valley. The claim is that a well designed humanoid robot approached in a relevant environment will not trigger a negative reaction. The second source of inspiration came from the model presented by Moore [8]. The expectation which may be formulated on the basis of this model is that for robot stimuli, which are less surprising for a subject the effect of the valley should be weaker. Thirdly, the intuitive approach to the uncanny valley hypothesis—especially for the case of computer generated models—suggests that such a model would appear more familiar and acceptable when presented in an expected and known context (e.g. from science fiction animations or computer games).

The paper is structured as follows. In the Sect. 2 we present methods and tools. We discuss the choice of models, experimental procedure and hypotheses. Section 3 presents results of the experiment. The last section covers discussion of the results, summary and future research plans. The paper is supplemented with an “Appendix”, where all the models used for the study are presented.

2 Methods

For our research we have prepared 12 computer-generated models. Models were retrieved from 3D character banks² and rendered with the use of the *Unity* environment.³ All the models were chosen arbitrarily by the authors and then con-

sulted with two designers experienced in game development. Models covered various levels of degrees of human-likeness (DOH)—ranging from simplistic robots, through androids, animated characters to human models. The 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 [6]. All the 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. This allows evaluation of body proportions, body elements (such as hands, feet) and even facial expressions.

Using these models we have prepared two sets of stimuli for the experiment: (A) models on a neutral background (empty room with uniform coloured greyish walls); (B) models on a background, which is suitable for a given model. By suitable we mean a background presenting a context which is relevant to a given model. Robots were placed in an industrial context of a factory, androids in a science-fiction scenery, animated characters on a cartoon-like background and humans in the backgrounds depicting a city. The main idea was that the model would be placed in a background where it would not be surprising to see it. The relevance of the background was assessed by the authors. It is worth to mention that models from one category were presented in the same surrounding, only the camera angle was different, so that the background consistency for each model category would be preserved (see the model set in the “Appendix”). An example of a model with two different backgrounds is presented in Fig. 1. All the backgrounds used for stimuli (A) and (B) had the same resolution (1920 × 1080 pixels). The same pose of a model was preserved for both sets of stimuli as only the background was modified change (models in group B do not interact with elements of a scene).

Neutral background (i.e. the room) was rendered with the use of the *Unity* environment. Different backgrounds for the group (B) were retrieved from <https://assetstore.unity.com/>.

2.1 Procedure

The experiment took place in the Reasoning Research Group Laboratory at the Institute of Psychology AMU. Stimuli were presented on a 22 inch computer screen with the use of Google Forms.⁴ It is worth to notice, that the reaction times were not measured during the experiment and there were no time constraints for our subjects.

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

³ <https://unity3d.com/>.

⁴ <https://forms.google.com>.

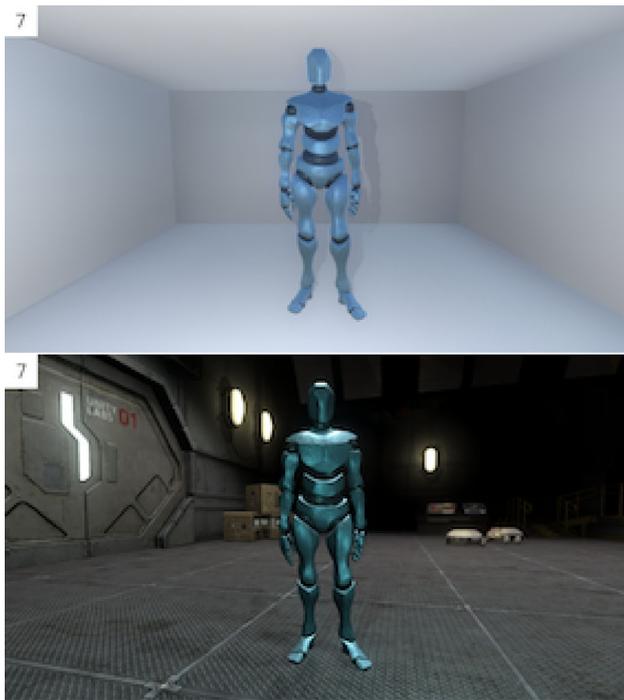


Fig. 1 Example of a model used in the study. Top picture presents the model on a neutral background (A) and the bottom picture presents the same model in a suitable background (B)

Subjects were divided into two groups. Group (A) was presented with models on a neutral background and group (B) with models on a suitable background.

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 these information, the subject pressed a key and started an experiment.

This phase consisted of presenting 12 models in a given variant (A—neutral background, B—suitable background). Each model was presented twice. First time with the following questions (below we present their translation, the experiment was run in Polish):

1. *How much does the presented model resembles 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. *How hard did you find answering the previous question?* (Answers on a scale 1–4, where (1) very easy, (2) easy, (3) hard, (4) very hard).

Second presentation of the model was accompanied by the following questions:

1. *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).
2. *What is your emotional reaction for the presented model?* (Three possible answers: the model makes me feel anxious, my reaction is neutral, I feel sympathy for the model. Answers were coded as: 0 for the emotional reaction means ‘neutral’, 1 is ‘sympathy’ and –1 is ‘anxiety’).

2.2 Subjects

64 subjects took part in the experiment. They were recruited among the cognitive science students at the Institute of Psychology AMU. They were randomly assigned to groups (A) and (B), accordingly to groups (A) and (B) of stimuli. The group (A) consisted of 32 subjects: 22 women and 10 men, mean age 20.69, $SD = 2.02$. The group (B) consisted of 32 subjects: 22 women and 10 men, mean age 20.72, $SD = 1.59$.

2.3 Hypothesis

The study addressed four aspects of the models:

1. degree of human-likeness,
2. difficulty of the human-likeness assessment,
3. comfort,
4. emotional reaction.

Our hypotheses were the following.

H1 Concerning the context condition:

H1a models presented on a neutral background will have lower comfort scores than the ones presented on a suitable background;

H1b different emotional reactions will be observed for models presented on a neutral background and the ones presented on a suitable background;

H2 is inspired by Kätsyri et al. [5]. The uncanny valley effect will be observed with respect to:

- comfort,
- human-likeness assessment difficulty,
- emotional reaction level.

3 Results

For the data analysis we used *R* statistical software ([11]; version 3.3.1). The summary of the results for group (A)—models on the neutral background and (B)—models with the

Table 1 The summary of the results for group (A)—models on the *neutral* background and (B)—models with the *suitable* background

Model	Human-likeness (median)		Difficulty (median)		Comfort (median)		Emotional reaction	
	A	B	A	B	A	B	A	B
1	2.0	2.0	2.0	2.0	2.0	2.0	0	0
2	5.0	5.0	1.0	1.0	2.0	2.0	0	0
3	3.5	3.0	2.0	2.0	3.0	3.0	-1	0
4	4.0	4.5	1.0	1.5	3.0	2.0	-1	0
5	2.0	2.0	2.0	2.0	3.0	3.0	-1	0
6	5.0	5.0	1.0	1.0	2.0	2.0	0	0
7	3.0	3.0	2.0	2.0	3.0	3.0	0	0
8	2.0	2.0	2.0	2.0	4.0	3.0	-1	-1
9	4.0	4.0	1.0	2.0	2.0	2.0	1	1
10	5.0	5.0	1.0	1.0	2.0	2.0	0	0
11	3.0	3.0	2.0	2.0	3.0	3.0	0	-1
12	4.0	4.0	2.0	2.0	3.0	2.5	0	0

Table 2 Mann–Whitney U test results for the *comfort of interaction* variable

Model	Result
1	W = 547.0 $p = 0.631$
2	W = 628.5 $p = 0.102$
3	W = 559.5 $p = 0.513$
4	W = 562.0 $p = 0.489$
5	W = 513.5 $p = 0.989$
6	W = 605.5 $p = 0.191$
7	W = 577.5 $p = 0.364$
8	W = 680.5 $p = 0.018^*$
9	W = 605.0 $p = 0.194$
10	W = 475.5 $p = 0.614$
11	W = 569.0 $p = 0.429$
12	W = 567.0 $p = 0.446$

* $P = 0.05$ **Table 3** Mann–Whitney U test results for the *emotional reaction* variable

Model	Result
1	W = 466.0 $p = 0.476$
2	W = 553.0 $p = 0.482$
3	W = 445.0 $p = 0.318$
4	W = 487.5 $p = 0.703$
5	W = 477.0 $p = 0.604$
6	W = 454.0 $p = 0.325$
7	W = 575.0 $p = 0.349$
8	W = 412.5 $p = 0.099$
9	W = 471.0 $p = 0.514$
10	W = 582.0 $p = 0.245$
11	W = 667.0 $p = 0.018^*$
12	W = 553.0 $p = 0.512$

* $P = 0.05$

suitable background is presented in Table 1. For the human-likeness, difficulty and comfort we provide median values. For the emotional reaction we provide the most frequent answer provided in a group.

3.1 Hypothesis 1

Let us start from the central hypothesis of this research, namely (**H1**) stating that a difference will be observed for the groups (A) and (B) with respect to comfort (**H1a**) and emotional reaction (**H1b**). As the data was not normally distributed, Mann-Whitney U test was used to study the differences between groups. The detailed results of the test are presented in Table 2 for the comfort of interaction and Table 3 for the emotional reaction.

Hypothesis (**H1a**) was not confirmed. Changes between models presented on a neutral and on a suitable background were observed for models 4, 8 and 12 (see medians presented in Table 1), however the only statistically significant difference was observed for model 8 (rated as ‘uncomfortable’ for the neutral background and as ‘neutral’ for the suitable background).

Hypothesis (**H1b**) was also not confirmed. Emotional reaction change may be observed for four models (see Table 1), however only for model 11 this change is significant (from ‘neutral’ into ‘anxiety’ when presented on a suitable background).

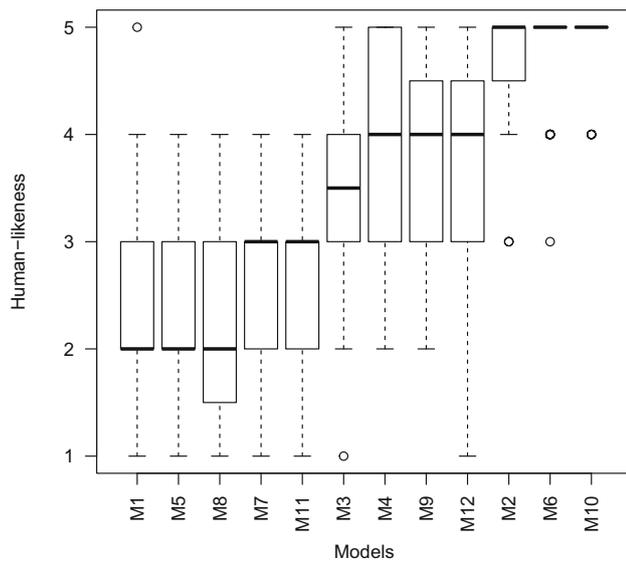


Fig. 2 Human-likeness assessment for group A. The ordering is accordingly to growing human-likeness. Numbers of models indicate the order in which they appeared in the study

3.2 Hypothesis 2

As hypothesis (H1a) and (H1b) were not confirmed, i.e. there are no significant differences between groups (A) and (B) for the further analysis we used only results from group (A). Let us now analyse the ordering of the models accordingly to DOH. In accordance with the results of our previous study with these models (see [6]) we have not assumed any pre-ordering of the presented models. The resulting DOH ordering is based solely on the subject's assessment. Figure 2 presents the ordering of the models accordingly to growing human-likeness. The resulting ordering is intuitive. On the left side of X axis we find robotic models, followed by animated characters and then with humans. Models are clearly grouped into three categories. It is worth to point out that animated characters (models 4, 9 and 12) were evaluated on the same level 4 (median), i.e. 'Rather human-like' and human models also received uniform scores at level 5 (median), i.e. 'Completely human-like'. For the robotic models we observe the variety from 2 ('Rather not human-like') to 3.5 (i.e., above 'It starts to look human-like').

Our second research hypothesis was inspired by Kätsyri et al. [5]. It states that the uncanny valley will be observed with respect to: comfort, human-likeness assessment difficulty, and emotional reaction level—i.e. that for our set of models we will observe more than one uncanny valley. Results with the respect to this hypothesis are presented in Fig. 3. Scales for comfort and difficulty are reversed (the higher the score is the harder it was to assess human-likeness or the more uncomfortable watching the model was), so the 'valley' is actually visible as a peak in the figure.

The effect is visible for model 12 for the comfort and difficult of DOH assessment. This 'valley' is with accordance with Kätsyri and colleagues descriptions, i.e. models on the left of 12 and on the right got better comfort and difficulty assessment scores.

When it comes to comfort, model 8 had the lowest score (median for the model is (4), i.e. 'uncomfortable'). This model resembles battle robots known from science-fiction films and computer games (see e.g. *Fallout*⁵, *Titanfall*⁶ or the *Halo* series⁷).

As for the emotional component, almost all models were assessed as 'neutral'—this applies to robotic models (1, 7, 11), animated model (12) and humans (2, 6, 10). Three robotic models (5, 8, 3) and one animated character (4) were evaluated as evoking anxious reactions. What is interesting, the 9th model (animated character) was the only one, which was evaluated as evoking 'sympathy'.

Hypothesis (H2) is therefore confirmed. The most apparent case is model 12, which evokes a 'valley' with respect to the comfort, difficulty of assessment and emotional reaction.

What is interesting is that the comfort dimension does not straightforwardly relate to the emotional relation component. This may be observed for model 8, as a comfort 'valley' is observed for it, but it was evaluated as 'neutral' when it comes for the emotional reaction.

4 Summary and Discussion

Hypothesis (H1) was not confirmed for the whole set of models. Statistically significant difference between two groups was observed only for single models.

When it comes to (H1a) it stated, that models presented on a neutral background would get worse results if we considered comfort dimension than the ones presented on a suitable background. This was observed for model 8, which is presented in Fig. 4.

A potential explanation for this fact may be the overall look of this model. It resembles a battle robot, which may trigger anxiety, when the model is presented in the empty room. When the same model is presented in a suitable background the comfort dimension improves, as it probably then more resembles characters know from science-fiction themed films or games.

As for hypothesis (H1b), we assumed that between groups differences with respect to the emotional reaction will be observed. This effect was not observed for all the models. One exception is model 11, for which the difference is statistically important. As for this fact it is more difficult to explain than

⁵ <https://www.fallout4.com/>.

⁶ <https://www.ea.com/pl-pl/games/titanfall/titanfall-2>.

⁷ <https://www.halowaypoint.com/en-us>.

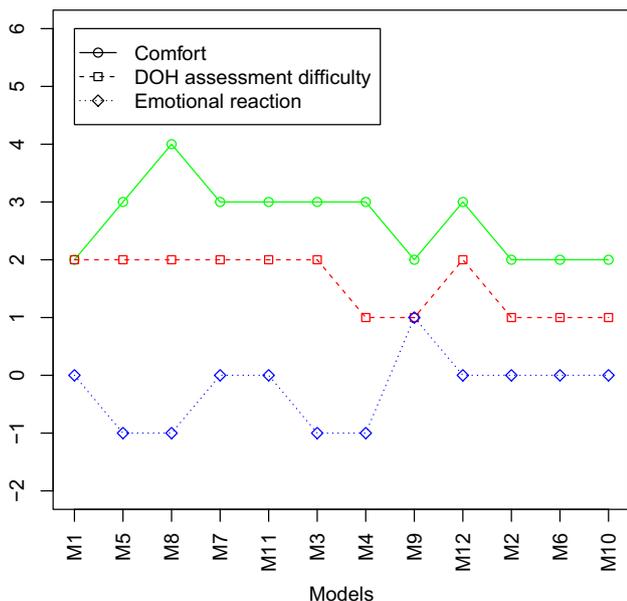


Fig. 3 Comfort, DOH assessment difficulty (medians) and emotional reaction (0 means ‘neutral’, 1 is ‘sympathy’ and –1 is ‘anxiety’) for group A. Models are ordered for growing human likeness. Numbers of models indicate the order in which they appeared in the study. Scales for comfort and difficulty are reversed, so the ‘valley’ is actually visible as a peak in the figure

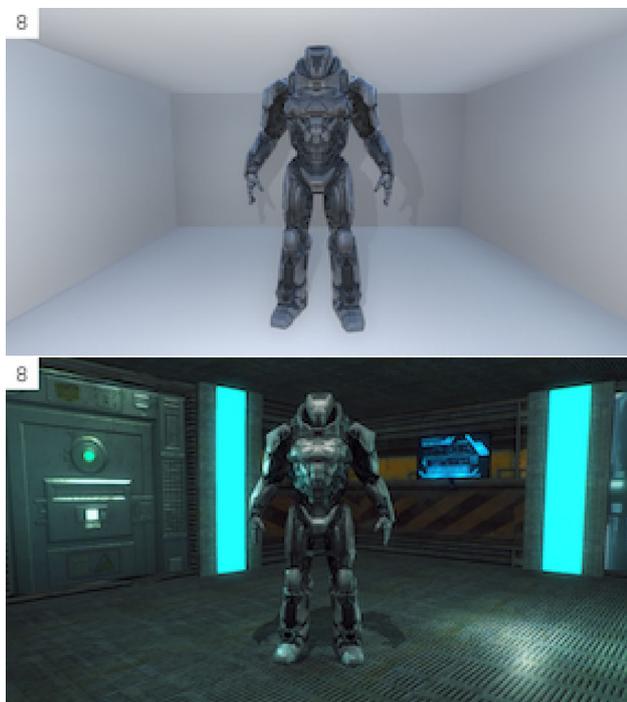


Fig. 4 Model 8 on two backgrounds

the one for model 8. When we analyse the overall look of the model, it shares the basic characteristics with model 7, for which the effect of emotional reaction difference was not observed (see Fig. 5).

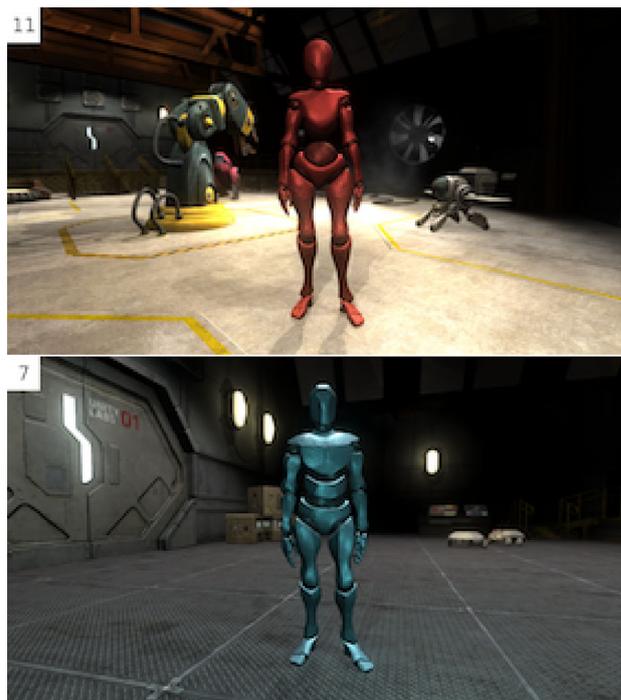


Fig. 5 Models 11 and 7 as they were presented on a suitable background. For model 11 statistically important difference in emotional reaction between groups were observed. What is interesting, such a difference was not observed for model 7

Accordingly to hypothesis (H2) many (independent) uncanny valleys should be observed for the presented models. This hypothesis was confirmed. For model 12 decrease in DOH difficulty assessment and comfort is clearly visible. There are two potential explanations for this fact. One is that possibly the discussed models (as animated character) fits to the Mori prediction. Its DOH reaches the level which triggers affinity. Other explanation, which we consider to be more probable, refers to the way this model is rendered. Lights on the scene are slightly unsuitable and make the model glow. As a result, this model differs from the other model from the research set. More research would be needed in order to evaluate this explanation.

Distinctive features (large head and eyes) may also explain the observed emotional reaction for model 9. It is the only model which was evaluated as triggering sympathy on an emotional reaction level. The facial expression of the model is also distinctive—it is the only model in our research with a visible smile. Face expression and its visibility will be a subject of our future research.

Our results are in line with predictions of [5]—many ‘valleys’ were observed. As such, this result of our study may serve as yet another confirmation that there is more into UVH studied than simply lowering affinity—which we find important as the number of UVH studies is growing, but we do not have a standardised experimental setting (see e.g. the dis-

cussion in [4]). However, even for this result there are still aspects that require further studies—at this point, we do not know whether and how they relate to each other. This issue certainly deserves further study. Especially collecting reports from the subjects containing explanations behind their evaluations would be helpful here.

It is also worth stressing that we are aware that certain results observed in our research may be due to the models we have used and/or rendered backgrounds. Models used for the experiment reported in this paper were previously tested in research described in [6]. On the basis of previous results we have resigned from zombie models in the stimuli set as they triggered strong emotional reactions (and thus for a somewhat obvious reason biased the comfort and DOH assessments). We have also resigned from the authoritative grouping of the models performed before the experiment (as it is often practised in the UVH research, see [1,10]). Instead we only used DOH assessment of our subjects. One of the unavoidable authoritative decisions in our study was however the choice of suitable backgrounds for group B. For future studies it would be recommended to use the competent judges method for evaluating possible backgrounds. What is more—as it was pointed out by an anonymous reviewer—future investigations are needed in case of views where the figure is a part of the scene (e.g., interacting with objects in the scene), or when there is no clear separation between figure and background.

Acknowledgements The authors would like to give their thanks to Marcin Jukiewicz, Petr Švarný and Mariusz Urbański for helpful feedback and comments on a draft of this paper.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interests.

Research Involving Human Participants The experiment was conducted in the laboratory conditions under a supervision of an exper-

imenter. Before the study all the subjects were informed about the aim of the study, its design and the rights of participants. On the basis of previous study [6] we have resigned from zombie models in the stimuli set as they triggered strong emotional reactions. Before the experiment, we have consulted the stimuli set used for the study with respect to this issue with the Reasoning Research Group member (during two seminar meetings).

Informed Consent The following information was read by the experimenter and displayed on the screen before the experiment started (we present translation, original information was in Polish).

We welcome you in the study devoted to the issues related to artificial intelligence. The study addresses the topic of artificial intelligence models and humans interaction.

During the study you will be presented with 12 models. For each model you will see 4 questions associated with a scale for providing answers. First question will address the issue of human-likeness of a given model. You will evaluate this on a scale 1 to 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.

The second question considers the issue how difficult was the evaluation required by the previous question. You will be asked to provide your answer on the following 1 to 4 scale, where: (1) very easy, (2) easy, (3) hard, (4) very hard

The third question will address the issue of comfort. How comfortable are you when watching a given model in a given environment. You will be asked to provide your answers on a scale 1–5, where: (1) very comfortable, (2) quite comfortable, (3) neutral, (4) uncomfortable, (5) very uncomfortable.

The fourth question will consider your emotional reaction to a given model. For this question you may choose from three possible answers: (1) the model makes me feel anxious, (2) my reaction is neutral, (3) I fell sympathy for the model.

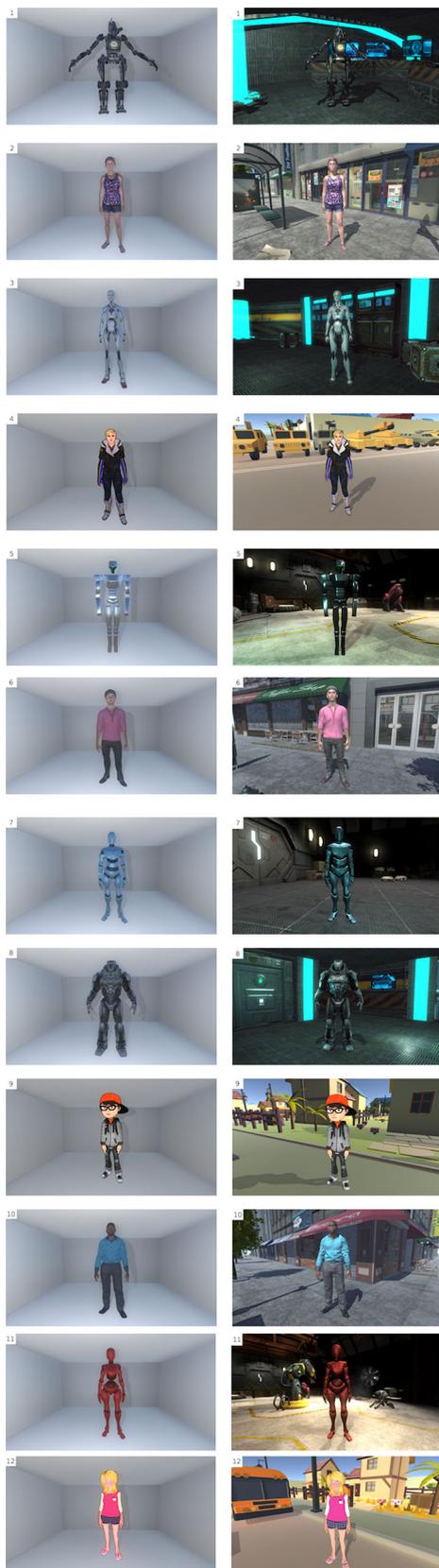
Hints and reminders will also appear also further in this study.

Please remember that the participation in this study is voluntary. You can resign from participation at any moment (even after it has started) without giving any reason.

The experiment results are anonymised and are processed only for the scientific purposes.

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.

Appendix



References

- Dill V, Flach LM, Hocevar R, Lykawka C, Musse SR, Pinho MS (2012) Evaluation of the uncanny valley in CG characters. In: Nakano Y, Neff M, Paiva A, Walker M (eds) Intelligent virtual agents. Springer, Berlin, pp 511–513
- Geller T (2008) Overcoming the uncanny valley. *IEEE Comput Graph Appl* 28(4):11–17
- Hanson D, Olney A, Prilliman S, Mathews E, Zielke M, Hammons D, Fernandez R, Stephanou H (2005) Upending the uncanny valley. In: Proceedings of the 20th national conference on artificial intelligence—volume 4, AAAI Press, AAAI'05, pp 1728–1729. <http://dl.acm.org/citation.cfm?id=1619566.1619636>
- Kätsyri J, Mäkäräinen M, Takala T (2017) Testing the ‘uncanny valley’ hypothesis in semirealistic computer-animated film characters: an empirical evaluation of natural film stimuli. *Int J Hum Comput Stud* 97:149–161
- Kätsyri J, Förger K, Mäkäräinen M, Takala T (2015) A review of empirical evidence on different uncanny valley hypotheses: support for perceptual mismatch as one road to the valley of eeriness. *Front Psychol* 6:390. <https://doi.org/10.3389/fpsyg.2015.00390>
- Łupkowski P, Rybka M, Dziedzic D, Włodarczyk W (2017) Human-likeness assessment for the uncanny valley hypothesis. *Bio Algorithms Med Syst* 13(3):125–131
- MacDorman KF, Green RD, Ho CC, Koch CT (2009) Too real for comfort? Uncanny responses to computer generated faces. *Comput Hum Behav* 25(3):695–710
- Moore RK (2012) A bayesian explanation of the ‘uncanny valley’ effect and related psychological phenomena. *Sci Rep* 2:864
- Mori M, MacDorman KF, Kageki N (2012) The uncanny valley [from the field]. *IEEE Robot Autom Mag* 19(2):98–100 (Original work published in 1970 in Japanese)
- Piwek L, McKay LS, Pollick FE (2014) Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition* 130(3):271–277
- R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>. Accessed 20.03.2017
- Ueyama Y (2015) A Bayesian model of the uncanny valley effect for explaining the effects of therapeutic robots in autism spectrum disorder. *PLoS ONE* 10(9):e0138,642

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Paweł Łupkowski Assistant Professor at the Department of Logic and Cognitive Science, Cognitive Science Curriculum Manager at the Institute of Psychology (Adam Mickiewicz University, Poznań). Co-founder of the Reasoning Research Group (<http://reasoning.edu.pl>). Research interests: philosophical logic (logic of questions), theories of reasoning, formal analysis of cognitive processes, serious games, games with a purpose, gamification in science and education.

Marek Rybka received MA in cognitive science from Adam Mickiewicz University. Reasoning Research Group member. His research interests cover human–computer interaction, human–robot interaction and usability.

Dagmara Dziedzic Currently, a Ph.D. candidate at Adam Mickiewicz University in Poznań. Her scientific interests are connected with the use of games in scientific research. A large part of her research is associated with the possibility to make games both useful tool for researchers and attractive entertainment for the player. Her work focuses on the transfer of mechanics known from traditional video games to games created for a scientific purposes. Her research is supported by her professional experience, since it is related to the design of commercial games and gamification systems.

Wojciech Włodarczyk focuses on designing and creating crowdsourcing systems used for acquiring linguistic resources. His scientific interests are also connected with creating artificial intelligence agents in computer games, as well as applying machine learning mechanisms in order to improve them.