"The Polar Express" is Bipolar: Critical Film Reviews Influence Uncanny Valley Phenomenon in Semi-Realistic Animation Films

Jari Kätsyri (jari.katsyri@maastrichtuniversity.nl)
Department of Cognitive Neuroscience, Oxfordlaan 55, 6229 EV Maastricht, The Netherlands

Abstract

Previous research suggests that semi-realistic animation films such as *The Polar Express* are representative of the uncanny valley (UV) hypothesis, which predicts that highly human-like artificial characters can appear eerie. In the present study, we investigated the extent to which critical film reviews can influence the perceived eeriness of such films. The reviews were adopted from authentic ones and expressed either negative or positive attitudes towards the animation techniques. Audiovisual speech asynchrony, which is known to induce eeriness, was included as an objective manipulation. Our results showed large review tone effects for both implicit and explicit eeriness evaluations. In contrast, speech asynchrony failed to elicit significant effects. These results demonstrate that critical film reviews representing opposite attitudinal poles can elicit consistent changes in the viewers' evaluations of semi-realistic animations. The present findings cannot, however, be taken as evidence against the UV hypothesis itself in computer-generated characters.

Keywords: Uncanny Valley hypothesis; anthropomorphism; social influence; animation films

Introduction

Realistic computer-generated (CG) characters are commonly used in lieu of human actors when creating social signal stimuli for neurocognitive experiments. Although this practice has several advantages such as the ease of creating and manipulating stimuli, it is conceivable that subtle flaws in highly realistic CG characters could elicit unintended negative reactions in human participants. The Uncanny Valley (UV) hypothesis predicted such reactions already in the 1970s. The original hypothesis suggested that observing highly realistic robots or other mechanical devices can elicit negative feelings characterized by eeriness and lack of familiarity (Mori, 1970/2012).

Empirical Evidence for Uncanny Valley

Although the UV hypothesis is compelling and it seems to have plentiful anecdotal evidence (see below), empirical evidence for its existence is still elusive (Kätsyri, Förger, Mäkäräinen, & Takala, 2015; S. Wang, Lilienfeld, & Rochat, 2015). A lack of commonly agreed-upon operationalization is a fundamental problem for testing the hypothesis. According to a conservative interpretation, any artificial-to-human stimulus continuum should elicit a non-linear evaluation curve in which the most negative evaluations occur at levels preceding the full degree of human-likeness. The bulk of empirical studies have, however, shown that increasingly human-like stimuli simply tend to elicit more positive evaluations in a linear manner (Kätsyri et al., 2015).

Although the UV is apparently not triggered by all artificial-to-human continua, it could still occur under some specific conditions. Accumulating evidence has shown that negative evaluations can be elicited by a mismatch between artificial and realistic features (Seyama & Nagayama, 2007; MacDorman & Chattopadhyay, 2016) or by categorization difficulty (Burleigh & Schoenherr, 2015; Yamada, Kawabe, & Ihaya, 2013; however, see MacDorman & Chattopadhyay, 2016). Almost all of this evidence comes from rigorous manipulations of CG and human faces. A few other studies using naturalistic – and plausibly more ecologically valid – stimuli have provided evidence for the UV in pictures of prosthetic hands (Poliakoff, Beach, Best, Howard, & Gowen, 2013) and real-world robot faces (Mathur & Reichling, 2016).

Uncanny Valley and Animated Film Characters

The UV hypothesis was rediscovered at the beginning of the present millennium (e.g., MacDorman, 2005) and has since received increasing research interest. First fully computer-animated films with deliberately realistic characters, including *Final Fantasy* (Aida, Lee, Sakai, Sakaguchi, & Sakakibara, 2001) and *The Polar Express* (Goetzman, Starkey, Teitler, & Zemeckis, 2004), were released roughly around the same time. Perhaps not coincidentally, these and some other animation films have been adopted as anecdotal examples of the UV, with frequent citations in scientific research reports (e.g., Piwek, McKay, & Pollick, 2014) and popular scientific magazines (e.g., Spinney, 2017).

It seems plausible that semi-realistic animated film characters could appear eerie in the sense of the UV. Early computer graphics methods in particular tended to suffer from shortcomings in modeling light reflections from the skin and eyes, for example (e.g., Wechsler, 2002). These and other subtle flaws on otherwise quite realistic characters could elicit sufficient featural mismatch to make them appear uncanny. In general, semi-realistic animated film characters stand out from the intentionally caricatured and exaggerated characters that are the norm in traditional animation (Kaba, 2012).

To our best knowledge, semi-realistic animated film characters, cartoonish animated film characters, and human actors have been explicitly compared with each other only recently in our own previous study (Kätsyri, Mäkäräinen, & Takala, 2017). The results showed that semi-realistic characters are considered more eerie and selected more often as representative examples of the UV hypothesis than cartoonish characters or human actors. These results suggest that semi-realistic animated film characters capture some aspects of the UV hypothesis.

Critical Film Reviews and Uncanny Valley

Semi-realistic animation films, which are a rare exception among the traditional computer-animated films, have attracted a mixed but predominantly negative critical reception. Some of the published reviews have explicitly characterized the realistic characters as "soulless" and "creepy" (e.g., Savlov, 2004). In the present study, we explore the extent to which this kind of critical reviews can influence the evaluation of animated characters in semi-realistic films.

A research tradition in social psychology has shown that individuals' attitudes and behavior are influenced by those of others (Bohner & Dickel, 2011). In one form of social influence, credible information from others is accepted as valid evidence about reality (Deutsch & Gerard, 1955). Perceived expertise on the subject matter has long been considered as one source of credibility (Kelman, 1956). On the other hand, negative attitudes are known to exert greater influence than equally intense positive attitudes (Cacioppo, Gardner, & Berntson, 1997). An excellent demonstration of this is that negative peer reviews exert greater influence than positive reviews on individuals' attitudes towards consumer products (Lee, Park, & Han, 2008). Taken together, negative reviews from film experts should exert an influence. Consistently, negative film reviews are known to elicit decreased box profit especially during early presentation weeks (Basuroy, Chatterjee, & Ravid, 2003).

Present Study

The present study aims to compare the effects of critical film reviews and objective differences between films on the UV phenomenon in semi-realistic animation films. When comparing small sets of individual films (as in Kätsyri et al., 2017), confound effects cannot be fully excluded. Instead, in the present study, we decided to manipulate audiovisual speech asynchrony in the same set of films. Audiovisual speech asynchrony was selected as the objective manipulation since a previous study has shown that it elicits eeriness in virtual characters (Tinwell, Grimshaw, & Nabi, 2015). Increased eeriness was observed at a 200-ms asynchrony at the earliest, and it was more pronounced when the auditory stream preceded rather than followed the visual stream. These findings are consistent with psychophysics literature, in which simultaneity judgments for audiovisual speech occur roughly in the time frame of –130 ms (audio first) to +220 ms (vision first) (Conrey & Pisoni, 2006).

Participants were asked both to rate the eeriness of films immediately after viewing them (implicit evaluations) and to score how representative each film was of UV after having received full debriefing (explicit evaluations). Our main hypotheses were:

H1: Majority negative reviews will elicit higher implicit and explicit eeriness than majority positive reviews.

H2: Asynchronous audiovisual speech will elicit higher implicit and explicit eeriness than synchronous audiovisual speech.

Method

Participants

Forty participants (20 women) with a mean age of 29.7 years (SD = 8.5) took part in the study. All participants were native Finnish speakers with a good command of written English. Three participants who scored poorly in a post-experimental reading comprehension test were replaced with new participants. Participants reported having normal hearing, normal (or corrected) visual acuity, and no history of dyslexia. Participants were compensated with two movie tickets. Study protocol was approved by the Aalto University Research Ethics committee.

Stimuli

Animation Film Scenes Stimuli were four film scenes from animation films *Final Fantasy* (Aida et al., 2001), *The Polar Express* (Goetzman et al., 2004)**,** *Beowulf* (Rapke, Starkey, & Zemeckis, 2007), and *Tintin* (Jackson, Kennedy, & Spielberg, 2011). Films were selected on the basis of our previous evaluation (Kätsyri et al., 2017). All films were fully computer-animated and used motion-capture techniques for character animation (most character motions were captured from real actors). Film scenes were extracted from official DVD releases, depicted spoken dialogue, and did not contain violence or nudity. These copyrighted materials were presented under an external license.

Audiovisual Speech Synchrony Original audiovisual tracks for the predominantly spoken film scenes were used for the synchronous speech condition. For the asynchronous condition, auditory stream in each film scene was modified to precede the visual stream by 200 ms.

Film Reviews For each film, six fictional but plausible film reviews (three positive and three negative) were created. For authenticity, all reviews were displayed in English. Available reviews were first extracted from Metacritic database (http://www.metacritic.com). Brief statements expressing positive and negative attitudes towards animation techniques were then extracted and modified to produce an initial set of 12 reviews (half positive). These reviews were evaluated by 19 additional participants for expressed attitude (−4 very negative to +4 very positive). Negative and positive reviews were clearly differentiated (M = −2.1 and +2.6, SD = 1.2 and 0.8), $F(1, 18) = 114.95$, $p < .001$, $\eta_p^2 = .87$.

Final selected stimuli were three positive and three negative reviews with roughly similar lengths (M = 82 and 79 words) and similar mean ratings across films. The reviews focused on either specific visual features (e.g., "[…] their eyes, supposedly the windows to the soul, are more often dead than alive") or the overall impression generated by highly realistic animations (e.g., "[…] the soft-edged, photorealistic style – suspended somewhere between live-action and animation, fairy tale and reality – feels entirely appropriate in this context").

Table 1: Counterbalancing between films and conditions.

Group	Negative/ Asynch.	Positive/ Asynch.	Negative/ Synch.	Positive/ Synch.
1	F1	F2	F3	F4
2	F2	F3	F4	F1
3	F3	F4	F1	F2
4	F4	F1	F2	F3

Procedure

Film preference evaluation was used as a distractor task to avoid making the true study objectives too obvious. Specifically, participants were explained that the study aimed to investigate the effects of critical film reviews on attitudes and consumer behavior. Full debriefing was given at the end of the study.

Participants were assigned randomly to the conditions of a 4 (group) × 2 (majority review tone: negative, positive) × 2 (speech synchrony: asynchronous, synchronous) mixed model design. Majority review tone and speech synchrony were counterbalanced with films using a 4 × 4 Latin square as shown in Table 1. Participants were asked to read film reviews carefully and told that a memory task would follow the evaluation. Each trial proceeded as follows:

- Title and a brief description of the film were presented with a minimum reading time of 20 s.
- Four reviews were presented (min. 10 s per each). For the majority negative/positive condition, three reviews were negative/positive and one was of the opposite valence. Reviews were paired randomly with fictional reviewer names (half female) and well-known magazine titles.
- An overview of the film scene was presented (min. 20 s).
- The movie scene was played back (164 s to 182 s).
- The participant answered a series of self-report questions.

The experiment was carried out on a desktop computer running Psychtoolbox (Brainard, 1997) for Matlab. Film stimuli were displayed on a 24" wide-screen display (Eizo ColorEdge CG241W) at horizontal resolution of 1024 pixels and vertical resolution depending on aspect ratio (424 to 548 pixels). Participants were seated 80 cm from the display. Auditory sound tracks were standardized at –5 dB and played on closed earphones at a loud but comfortable level. The whole experiment took approximately 60-90 mins.

Dependent Variables

Film Preference For attitude towards the film (cf. Voorveld, 2011), participants rated whether they enjoyed each film, were content with it, and found it interesting. This scale had a good internal reliability (Cronbach's $\alpha = 0.96$). For consumption intent (cf. A. Wang, 2006)**,** participants rated whether they would like to see the full film (again if seen), recommend it to a friend, and pay for seeing it (good reliability: $\alpha = 0.91$). These and all other ratings were given on a 7-step Likert scale ranging from total disagreement (1) to total agreement (7).

Table 2: Means (and SEMs) by majority review tone.

Variable	Negative	Positive	$F(1, 36)$	p	η_p^2
Film attitude	4.6 (0.2)	5.0 (0.1)	6.94	.027	.13
Consumption intent	3.8 (0.2)	4.4 (0.2)	8.72	.006	.19
Human-likeness	4.7 (0.2)	5.0 (0.1)	8.08	.007	.18
Eeriness	3.2 (0.2)	2.8 (0.2)	7.46	.010	.17
Representativeness	10.5 (0.7)	8.2 (0.8)	6.52	.015	.15

Eeriness (Implicit) and Human-likeness For eeriness, participants rated whether the characters appeared eerie, creepy, and strange (good reliability: $\alpha = 0.84$). Although not included in any hypothesis, human-likeness scale was also included because it is a focal dimension of the UV (Mori, 1970/2012). Participants were asked to rate whether the characters appeared realistic, cartoonish, and similar to real people (adequate reliability: $\alpha = 0.71$).

Representativeness (Explicit Eeriness) After finishing the film evaluation task, participants received a full debriefing of the experiment and the UV hypothesis. Participants were then given four plastic cards depicting the four animation films and asked to place these cards on a cardboard depending on how representative the films were of the UV (left side: not at all, right side: perfectly). Responses were scored from 0 to 100 based on the cards' physical positions.

Analysis

Data were subjected to a mixed-design GLM analysis in SPSS (version 24). Effect sizes were quantified using partial η^2 values and classified as large ($\eta^2 \geq .14$), medium ($\eta^2 \geq .06$), or small ($\eta^2 \geq .01$) based on Cohen's (1992) guidelines.

Results

Manipulation Checks

After the experiment, participants were asked to read 12 review statements for each film, and to tell whether they had read these reviews during the experiment (yes/no). Statements included four correct and eight incorrect options and an equal number of positive and negative reviews. Recognition performance, as indexed by *d'* sensitivity score (Stevens & Pashler, 2002), was well above chance level ($M = 1.47$, $SD = 0.59$), $t(39) = 15.80$, $p < .001$, showing that participants had attended and comprehended the reviews.

Negative Reviews

As can be seen in Table 2, majority review tone elicited significant effects for all dependent variables. Supporting H1, majority negative reviews elicited higher eeriness ratings and representativeness scores than majority positive reviews. In the distractor task, majority negative reviews elicited decreased film attitude and consumption intent ratings. Human-likeness ratings were decreased for majority negative review reviews. Review tone had a medium effect for film attitude and a large effect for all other variables.

Table 3: Means (and SEMs) by speech synchrony.

Variable	Asynch.	Synch.	$F(1, 36)$	p	η_p^2
Film attitude	4.8 (0.2)	4.8 (0.2)	0.02	.884	.00
Consumption intent	4.2 (0.2)	4.0 (0.2)	0.46	.503	.01
Human-likeness	4.8 (0.2)	4.9 (0.2)	0.21	.650	.01
Eeriness	3.0 (0.2)	3.0 (0.2)	0.07	.790	.00
Representativeness	9.9 (0.8)	8.7 (0.9)	1.26	.269	.03

Speech Asynchrony

As shown in Table 3, speech synchrony failed to elicit significant main effects on eeriness, representativeness, or any other variable. Asynchrony elicited slightly higher representativeness scores than synchrony; however, this effect was small and clearly non-significant. Hence, the results did not support hypothesis H2.

Interactions

Fig. 1 illustrates interaction effects between review tone and speech synchrony. Visual inspection suggests that review effects for some variables were qualified by an interaction with speech synchrony. Consistently, significant and large interaction effects were observed for film attitude, $F(1, 36) = 7.05$, $p = .012$, $\eta_p^2 = .16$, consumption intent, $F(1, 36) = 6.42$, $p = .016$, $\eta_p^2 = .15$, and human-likeness, $F(1, 36) = 7.59$, $p = .009$, $\eta_p^2 = .17$. Simple effect tests revealed that for these variables, negative reviews elicited lower ratings than positive reviews only in the asynchronous condition ($p \leq .001$, $\eta_p^2 \geq .254$). Importantly, interaction effects were not significant for eeriness, $F(1, 36) < 1$, $p = .643$, $\eta_p^2 = .01$, or representativeness, $F(1, 36) = 1.21$, $p = .278$, $\eta_p^2 = .03$. Hence, H1 was not affected by this interaction.

Confounds

The following confound effects were tested and excluded using mixed model analyses (Hoffman & Rovine, 2007): previous familiarity (film seen or not seen), awareness of UV, awareness of review manipulation, and awareness of asynchrony manipulation. A mixed model equivalent to GLM was first specified, and confounds were then added and tested individually. The pattern of significant results was not changed by the inclusion of any tested confound into the model.

Awareness variables were derived from a semi-structured interview conducted the end of the experiment. Briefly, UV awareness meant that participants (18%) had heard about the UV hypothesis, were able to explain it correctly, and associated it with this experiment. For review awareness, participants (30%) noticed that films were preceded unequally by positive and negative reviews, were able to choose which films received majority positive/negative reviews, and were aware of the manipulation. For asynchrony awareness, participants (30%) mentioned audiovisual asynchrony spontaneously and were aware of the manipulation. Notably, only 8% of participants were able to choose asynchronous films correctly, however.

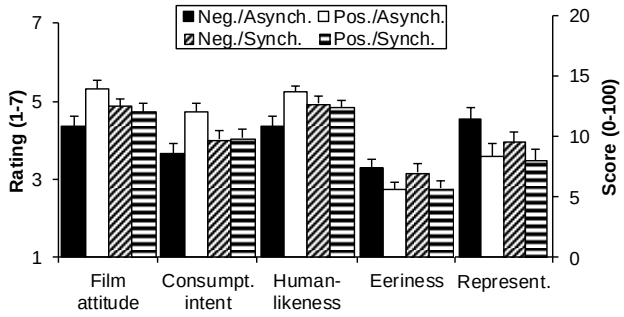


Figure 1: Means (and SEMs) by majority review tone and speech synchrony. Note that representativeness scores are truncated to 0–20.

Discussion

The aim of the present study was to provide an objective comparison of the effects of critical reviews and objective differences (audiovisual speech synchrony or asynchrony) on the UV phenomenon in semi-realistic animation films. We hypothesized that critical film reviews focusing on the appearance of realistic animation technologies would exert an influence on both implicitly and explicitly evaluated eeriness. Indeed, our results showed that review tone (majority positive or negative) elicited large effects both on eeriness ratings given immediately after film viewing (implicit evaluations) and representativeness scores given after full debriefing of the UV (explicit evaluations). Unexpectedly, audiovisual speech asynchrony failed to elicit significant effects.

The present results demonstrate that social influence (Bohner & Dickel, 2011) originating from critical reviews can exert a large influence on the subjective eeriness of semi-realistic animated film characters. Results from explicit evaluations are particularly important because they demonstrate that critical reviews can affect individuals' attitudes specifically in the UV context. A closer inspection of the results suggests that participants did not consider the animation films as particularly eerie: 95% confidence interval for the overall eeriness ratings was [2.7, 3.3] on the 7-step scale (i.e., below midpoint) and for the overall representativeness scores [8.1, 10.5] on the visual scale from 0 to 100. In the absence of strong effects in the stimuli, participants may have relied on information from the film reviews instead.

Although not predicted beforehand, majority negative reviews also elicited lower human-likeness ratings than majority positive reviews – but only in film scenes with asynchronous audiovisual speech dialogue. This result differs clearly from eeriness evaluations, which were not sensitive to the interaction between review tone and synchrony. This suggests that review effects were more robust for evaluated eeriness than for evaluated human-likeness.

The failure to find significant effects for the 200-ms audiovisual speech asynchrony is surprising given that such delay should be noticeable on human faces (Conrey & Pisoni, 2006) and elicit eeriness in virtual characters as well

(Tinwell et al., 2015). We note that several participants (30%) commented spontaneously upon asynchrony in the films even though much fewer (8%) were actually able to identify which films were truly asynchronous. A possible explanation for this discrepancy is that animated lip movements were sufficiently inaccurate to begin with in some of the films so that additional delay had little or no effect on the perceived asynchrony. It is also possible that asynchrony effects may have been concealed by the stronger review tone effects. In partial support, secondary evaluations – but not eeriness evaluations – were affected by the interaction between review tone and asynchrony.

A significant limitation of the present study is that, in the absence of statistically significant effect for asynchronous speech, the magnitude of critical review effect cannot be meaningfully interpreted. Hence, the present results can tell us that critical film reviews exert a large influence on the perceived eeriness of semi-realistic animated film characters, but it cannot tell us whether these effects are weaker or stronger than those elicited by genuine physical differences in the stimuli.

The present results are nevertheless relevant for the UV phenomenon because semi-realistic animated film characters have already been linked to this phenomenon previously (Kätsyri et al., 2017). Methodological limitations in this previous study warrant some caution in interpreting the findings; for example, eeriness ratings were close to minimum for all film characters. It is particularly clear that semi-realistic animation films do not elicit any such aversion that could be implicitly or explicitly likened to human corpses or zombies (cf. Mori, 1970/2012). Nevertheless, semi-realistic animated characters received slightly but statistically significantly higher eeriness ratings and were considered more often as being representative of the UV hypothesis than other types of films. Hence, these previous findings provide support for subtle UV effects in semi-realistic animated film characters. The present results add to this by demonstrating that such effects can also be elicited by critical film reviews. It should be emphasized that the present study did not aim to "find" the UV for animated characters, as this was already done previously.

Precisely for the same reason, the present results should be considered meaningful even though the review effect was not tested for cartoonish animated characters or real human actors. In fact, the present film reviews that focused explicitly on the disadvantages of realistic animation techniques could not have been paired with any other types of films, at least not without arousing suspicion in the study participants. Although fictitious negative reviews focusing on the disadvantages of traditional animation could possibly have been created, it is likely that such stimuli would have been considered implausible as this kind of reviews do not appear to exist in reality.

Taken together, the present findings demonstrate that critical film reviews representing opposite attitudinal poles – negative or positive – towards realistic animation techniques can elicit consistent changes in individuals' evaluations. In this sense, the UV phenomenon in semi-realistic animation films could be characterized as being bipolar.

The present findings should obviously be interpreted with caution when it comes to explaining the complexities of attitude formation in real life. Allowing some speculation, it is possible that critical film reviews might have contributed to the wide-spread adoption of specific animation films as anecdotal examples of the UV. The present results empathetically cannot be taken to mean that the UV would be just a media phenomenon, however. First, the present study does not allow comparing the effects of critical film reviews to genuine differences between stimuli. Second, empirical evidence for the UV phenomenon has already begun to accumulate, especially from studies with featurally mismatching (Seyama & Nagayama, 2007; MacDorman & Chattopadhyay, 2016) and naturalistic stimuli (Mathur & Reichling, 2016; Poliakoff et al., 2013).

Overall, the present findings highlight the importance of social factors in evaluating contemporary technological artefacts, in particular those that involve human-like characteristics.

Acknowledgments

This study was carried out at the Department of Computer Science, Aalto University, Finland, with an individual research grant from the Emil Aaltonen Foundation. The work was also supported by the Marie Skłodowska-Curie Individual Fellowship (H2020 programme) under grant agreement 703493. The author wishes to thank Prof. Tapio Takala, M.Sc.(Tech.) Meeri Mäkäräinen, and D.Tech. Klaus Förger for previous collaboration, which provided the foundation for the present study.

References

Aida, J., Lee, C., Sakai, A. (Producer), Sakaguchi, H., & Sakakibara, M. (Director). (2001). *Final Fantasy: The Spirits Within [Motion Picture]*. USA, Japan: Columbia Pictures.

Basuroy, S., Chatterjee, S., & Ravid, S. A. (2003). How critical are critical reviews? The box office effects of film critics, star power, and budgets. *Journal of Marketing*, *67*(4), 103–117. doi:10.1509/jmkg.67.4.103.18692

Bohner, G., & Dickel, N. (2011). Attitudes and attitude change. *Annual Review of Psychology*, *62*(1), 391–417. doi:10.1146/annurev.psych.121208.131609

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.

Burleigh, T. J., & Schoenherr, J. R. (2015). A reappraisal of the uncanny valley: categorical perception or frequency-based sensitization? *Cognitive Science*, *5*, 1488. doi:10.3389/fpsyg.2014.01488

Cacioppo, J. T., Gardner, W. L., & Berntson, G. G. (1997). Beyond bipolar conceptualizations and measures: The case of attitudes and evaluative space. *Personality and Social Psychology Review*, *1*(1), 3–25.

Cohen, J. (1992). A power primer. *Quantitative Methods in Psychology*, *112*(1), 155–159.

Conrey, B., & Pisoni, D. B. (2006). Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *The Journal of the Acoustical Society of America*, *119*(6), 4065–4073.

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology*, *51*(3), 629.

Goetzman, S., Starkey, S., Teitler, W. (Producer), & Zemeckis, R. (Director). (2004). *The Polar Express [Motion Picture]*. USA: Warner Brothers Pictures.

Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: Foundations and illustrative examples. *Behavior Research Methods*, *39*(1), 101–117. doi:10.3758/BF03192848

Jackson, P., Kennedy, K. (Producer), & Spielberg, S. (Director). (2011). *The Adventures of Tintin: The Secret of the Unicorn [Motion Picture]*. USA: Columbia Pictures.

Kaba, F. (2013). Hyper-realistic characters and the existence of the Uncanny Valley in animation films. *International Review of Social Sciences and Humanities*, *4*(2), 188–195.

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. *Frontiers in Psychology*, *6*, 390. doi:10.3389/fpsyg.2015.00390

Kätsyri, J., Mäkäräinen, M., & Takala, T. (2017). Testing the "uncanny valley" hypothesis in computer-animated film characters: An empirical evaluation of natural film stimuli. *International Journal of Human-Computer Studies*, 97, 149–161.

Kelman, H. C. (1956). Processes of opinion change. *Public Opinion Quarterly*, *25*, 57–78.

Lee, J., Park, D.-H., & Han, I. (2008). The effect of negative online consumer reviews on product attitude: An information processing view. *Electronic Commerce Research and Applications*, *7*(3), 341–352. doi:10.1016/j.elerap.2007.05.004

MacDorman, K. F. (2005). Androids as an experimental apparatus: Why is there an uncanny valley and can we exploit it? (pp. 108–118). Presented at the CogSci-2005 Workshop: Toward Social Mechanisms of Android Science, Stresa, Italy.

MacDorman, K. F., & Chattopadhyay, D. (2016). Reducing consistency in human realism increases the uncanny valley effect; increasing category uncertainty does not. *Cognition*, *146*, 190–205. doi:10.1016/j.cognition.2015.09.019

Mathur, M. B., & Reichling, D. B. (2016). Navigating a social world with robot partners: A quantitative cartography of the Uncanny Valley. *Cognition*, *146*, 22–32. doi:10.1016/j.cognition.2015.09.008

Mori, M. (1970/2012). The Uncanny Valley (K. F. MacDorman & N. Kageki, Trans.) *IEEE Robotics & Automation Magazine*, *19*(2), 98–100. doi:10.1109/MRA.2012.2192811

Piwek, L., McKay, L. S., & Pollick, F. E. (2014). Empirical evaluation of the uncanny valley hypothesis fails to confirm the predicted effect of motion. *Cognition*, *130*(3), 271–277. doi:10.1016/j.cognition.2013.11.001

Poliakoff, E., Beach, N., Best, R., Howard, T., & Gowen, E. (2013). Can looking at a hand make your skin crawl? Peering into the uncanny valley for hands. *Perception*, *42*(9), 998–1000. doi:10.1068/p7569

Rapke, J., Starkey, S. (Producer), & Zemeckis, R. (Director). (2007). *Beowulf [Motion Picture]*. USA: Paramount Pictures.

Savlov, M. (2004, November 12). The Polar Express. *The Austin Chronicle*. Retrieved from http://www.austinchronicle.com/calendar/film/2004-11-10/the-polar-express/

Seyama, J., & Nagayama, R. S. (2007). The uncanny valley: Effect of realism on the impression of artificial human faces. *Presence: Teleoperators and Virtual Environments*, *16*(4), 337–351. doi:10.1162/pres.16.4.337

Spinney, L. (2017, October 29). Exploring the uncanny valley: Why almost-human is creepy. *New Scientist*.

Stevens, S. S., & Pashler, H. E. (Eds.). (2002). *Stevens' Handbook of Experimental Psychology* (3rd ed). New York: John Wiley & Sons.

Tinwell, A., Grimshaw, M., & Nabi, D. A. (2015). The effect of onset asynchrony in audio visual speech and the uncanny valley in virtual characters. *International Journal of the Digital Human*, *2*(2), 97–110.

Voorveld, H. A. M. (2011). Media multitasking and the effectiveness of combining online and radio advertising. *Computers in Human Behavior*, *27*(6), 2200–2206. doi:10.1016/j.chb.2011.06.016

Wang, A. (2006). The effects of expert and consumer endorsements on audience response. *Journal of Advertising Research*, *45*(04), 402. doi:10.1017/S0021849905050452

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The uncanny valley: Existence and explanations. *Review of General Psychology*, *19*(4), 393–407. doi:10.1037/gpr0000056

Wechsler, L. (2002, January 6). Why is this man smiling? *Wired*, 10.06.

Yamada, Y., Kawabe, T., & Ihaya, K. (2013). Categorization difficulty is associated with negative evaluation in the "uncanny valley" phenomenon. *Japanese Psychological Research*, *55*(1), 20–32. doi:10.1111/j.1468-5884.2012.00538.x