doi:10.1068/p6492

Crossing the 'uncanny valley': adaptation to cartoon faces can influence perception of human faces

Haiwen Chen¶, Richard Russell§, Ken Nakayama

Department of Psychology, Harvard College, Harvard University, Cambridge, MA 02138, USA; ¶ and Functional Neuroimaging Laboratory, Brigham and Women's Hospital, 824 Boylston Street, Chestnut Hill, MA 02467, USA; § and Department of Psychology, Gettysburg College, Gettysburg, PA 17325, USA; e-mail: haiwen.chen@post.harvard.edu

Margaret Livingstone

Department of Neurobiology, Harvard Medical School, Harvard University, Boston, MA 02115, USA Received 19 June 2009, in revised form 23 December 2009; published online 22 March 2010

Abstract. In this study we assessed whether there is a single face space common to both human and cartoon faces by testing whether adaptation to cartoon faces can affect perception of human faces. Participants were shown Japanese animation cartoon videos containing faces with abnormally large eyes. The use of animated videos eliminated the possibility of position-dependent retinotopic adaptation (because the faces appear at many different locations) and more closely simulated naturalistic exposure. Adaptation to cartoon faces with large eyes significantly shifted preferences for human faces toward larger eyes, consistent with a common, non-retinotopic representation for both cartoon and human faces. This supports the possibility that there are representations that are specific to faces yet common to all kinds of faces.

1 Introduction

Personal beauty affects many facets of our lives. We attribute a wide array of positive qualities, such as being more occupationally and interpersonally competent, better adjusted, and having greater social appeal, to attractive people and their opposites to unattractive people (Dion et al 1972; Feingold 1992; Langlois et al 2000). Given the importance of these judgments, we might ask how we make judgments on attractiveness.

Some aspects of attractiveness appear to be universal. Cross-cultural studies show that individuals from different cultures agree on which faces in a set are more attractive (Bernstein et al 1982; Cunningham et al 1995; Geldart et al 1999; Perrett et al 1994). Furthermore, developmental studies show that young babies look longer at faces rated as attractive by adults than at those rated as unattractive (Geldart et al 1999; Langlois et al 1993; Samuels et al 1994; Slater et al 1998).

Although there is extensive evidence for universal factors of attractiveness, there are also systematic differences in preferences between individuals, which might reflect differences in experience. Martin (1964) found that, while both African-American and Caucasian men considered Caucasian features to be more attractive than African features in the USA, the same evaluation conducted in Nigeria showed a preference for African features. Furthermore, Bronstad and Russell (2007) suggested that pairs with close relationships, and presumably more shared experiences, demonstrate even greater agreement on facial attractiveness than strangers from the same race or culture.

What individuals identify as an attractive face has also been extensively shown to shift with adaptation in laboratory studies. Peskin and Newell (2004) demonstrated that increasing exposure and thereby increasing the familiarity of faces increased their attractiveness. More specifically, Rhodes et al (2003) showed that adaptation to consistent facial distortions shifted the most attractive face toward the distortion. For instance, after repeated exposure to faces with wide noses, subjects reported faces with wider noses as more attractive. Cooper and Maurer (2008) showed similar effects with the vertical position of features on a face.

The construct of 'face space' helps to explain these adaptive shifts. If representations of faces are organized in a multidimensional 'face space' centered on a prototypical face that represents the mean of a person's experience, faces are judged by proximity to the prototype (Valentine 1991). Adaptation to a consistent distortion has been suggested to shift the prototype toward the adapting distortion (Leopold et al 2001, 2005; Rhodes et al 2003, 2004b; Webster et al 2004). Neuroimaging evidence supports this view, finding increased signal from the fusiform face area, an area proposed to be specialized for face processing, with increased variation from mean faces (Loffler et al 2005). Moreover, a shift in neural tuning was proposed to produce a shift in the position of the mean face. Attractiveness judgments have been suggested to parallel this shift of the mean face (Rhodes et al 2003).

However, it is assumed that these studies done in controlled laboratory settings in fact reflect real-world exposure and explain real-world differences in preferences for faces. We test the validity of this assumption in this study by using a more naturalistic but still controlled tool for adaptation: videos. This offers a more naturalistic exposure because the position and size of the adapting faces change dynamically, because accompanying speed and movement enhance life-like qualities, and because watching videos/television itself is common behavior.

Furthermore, in this study we use an adaptation paradigm to assess the scope of face space. Past work found common and group-selective mechanisms for coding faces of different races and different genders, suggesting that there is a common coding mechanism for human faces (a single face space) and dissociable coding mechanisms for subgroups of human faces (Jaquet and Rhodes 2008; Jaquet et al 2007a, 2007b; Little et al 2005, 2008; Rhodes et al 2003, 2004a, 2007; Webster et al 2004). But does this face space extend to other kinds of faces, cartoon faces for instance?

While it might seem obvious that this general face representation would extend to cartoon faces since cartoons resemble humans, the concept of the 'uncanny valley' proposed by Mori (1970), suggests this may not be the case. Mori proposed a curve to show the relationship between peoples' positive responses to an object and the humanness of that object. He suggested that as human-like features increase, people respond more positively to the object. However, at a distinct point, the curve dramatically dips into the 'uncanny valley', a region where the deviations from humanness are stronger than the reminders of humanness, so a feeling of uncanniness overcomes familiarity, and people respond with repulsion (figure 1). This effect has been supported by controlled empirical studies (MacDorman et al 2009; Seyama and Nagayama 2007; Steckenfinger and Ghazanfar 2009). Specifically, MacDorman et al (2009) found that faces were especially prone to eeriness when photorealism of eyes versus photorealism of skin were mismatched, creating conflicting degrees of human likeness. This points to a divide between human faces and other groups of faces, and suggests that there may not be a common, fundamental representation of faces.

Thus, we assess whether there is a common coding mechanism for all faces (a general face space), including cartoon faces, by testing if adaptation can have effects across categories of face, from cartoon to human. It is important to note that there are different mechanisms for visual adaptation. Xu et al (2008) found that adaptation to concave or convex curves led subjects to perceive faces as happy or sad, respectively, more frequently. However, the curves had to be positioned in the same spatial location as the mouths of the faces for this effect to occur, meaning that the effect is a result of low-level retinotopic adaptation. The fact that faces in videos appear in different locations ensures that adaptation effects cannot be primarily attributed to low-level retinotopic adaptation but instead would indicate higher-level adaptation.

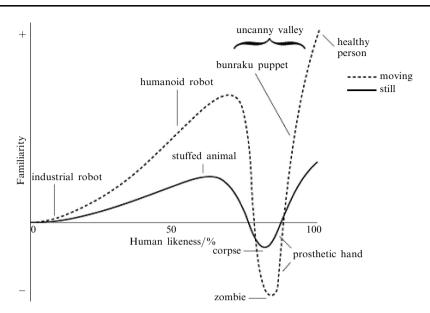


Figure 1. Mori's 'uncanny valley'. Familiarity and positive feelings increase with increasing human likeness until it reaches the uncanny valley where the imperfections cause a feeling of strangeness and repulsion. As we approach a normal human, the positive feelings rise again.

Participants were exposed to cartoon videos containing characters with very large eyes. Before and after the adaptation phase, we presented human faces and asked participants to rate them for attractiveness. If adaptation to cartoon faces does in fact affect attractiveness ratings for human faces, we expect adaptation to shift individual preferences toward the distortion such that participants will rate faces with larger eyes as more attractive after adaptation than before adaptation. This would suggest that adaptation in a laboratory setting does in fact generalize to a more naturalistic setting to explain real-world differences in preferences. This would also suggest that we have a basic representation of a face general to all kinds of faces.

2 Methods

2.1 Participants

Participants were forty-five Harvard students, fourteen males and thirty-one females, between the ages of 17 and 25 years (mean, M = 19.26 years, SD = 1.61 years) who consented to participating in the study and self-identified as having had no exposure to Japanese animation before participating in the study. We included only individuals with no prior exposure in order to ensure a more homogeneous background experience amongst subjects to the adaptation stimuli. Participant ethnicity included twenty-five Caucasian, nine Asian, six African/African-American, four Hispanic, and one Mediterranean individuals. All were recruited from Harvard University through the study pool and received course credit for participation.

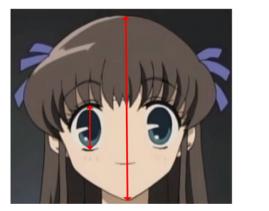
Participants were assigned to two groups generally matched for gender and ethnicity (not for age because of the small range); one group was exposed to cartoons, and the other was exposed to a live-action video with real actors.

2.2 Stimuli

2.2.1 Adaptation stimuli. The test group watched a 50-min cartoon video—the first two episodes of the Japanese animation series *Fruits Basket*. The control group viewed a 50-min live action video with real actors—the first episode of the television series *Heroes*.

Both videos included presentation of both male and female faces for almost the entirety of the 50 minutes. Both shows presented 7 female faces, not including background extras; *Heroes* presented 12 male faces and *Fruits Basket* presented 5. In terms of ethnicity, *Heroes* presented actors of varying ethnicities, and the cartoon characters in *Fruits Basket* were of undefined ethnicity, having various shades of hair and eye color and no distinct ethnic features. Both shows have a storyline about the lives of individuals with supernatural abilities that are both gifts and curses, though *Heroes* has a darker emotional tone than *Fruits Basket*.

To compare relative eye size between videos, the ratio of the length and width of the eyes to those of the head were measured from screenshots (see figure 2). Three of the most viewed characters of each gender from both of the shows were used for comparison. For females, the average eye-to-head length ratio for the animated characters in *Fruits Basket* (M = 0.18) was almost four times greater than the ratio for the actors in *Heroes* (M = 0.05). The average width ratio was essentially the same (M = 0.21 for animated, M = 0.23 for live action). For males, the average length ratio for animated characters (M = 0.11) was almost three times greater than that of the actors (M = 0.04). As with females, the male width ratio was comparable (M = 0.31 for animated, M = 0.21 for live action).



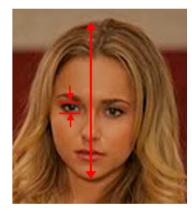


Figure 2. [In colour online, see http://dx.doi.org/10.1068/p6492] Screenshots of one of the main female characters from the animated show and the live action show, representing characters of comparable age. The lines represent measurements used to approximate the ratios of eye-to-head length. The cartoon face has a ratio about four times that of the human face.

2.2.2 Test stimuli. Participants were asked to rate for attractiveness a series of human faces with different eye sizes before and after adaptation. The set of stimuli consisted of 120 images of female faces viewed from the front with neutral expressions in identical conditions. 60 of the images were faces with different eye sizes created from 6 photographs of different female faces of Asian and Caucasian race/ethnicity. For each of the 6 distinct faces, 10 face stimuli were created with systematic distortion in the size of the eyes (see figure 3). The original photographs were drawn from the set in Russell (2009). They were altered with Photoshop to yield two different images, one with eyes that were 80% of the original size and one with eyes that were 120% of the original size. The two images were then morphed together in increments to yield 10 images of each of the faces with systematically different eye sizes. While the cartoon eyes are enlarged primarily vertically, the distortions of the test stimuli were made in both vertical and horizontal dimensions in order to maintain a natural appearance. All other aspects of the photographs were kept constant. The original photographs were altered versions.

80%

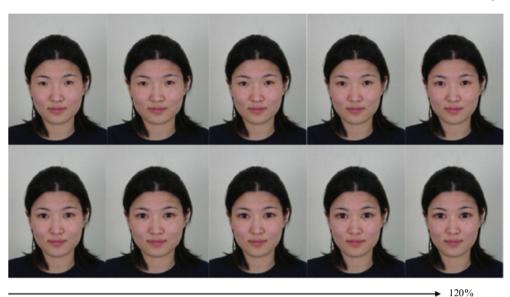


Figure 3. [In colour online.] An example of a set of stimuli with systematically varied eye size. They were created by incrementally morphing together the face with eyes 80% (upper left) of the original size and the face with eyes 120% (bottom right) of the original size. Participants viewed full colour images.

60 other images from the set of photographs in Russell (2009) were included to obscure the purpose of the study. The faces were shown in 10 blocks of the following pattern: face 1 eye variation, other face, face 2 eye variation, other face, face 3 eye variation, other face, face 4 eye variation, other face, face 5 eye variation, other face, face 6 eye variation, other face. The specific eye variation in each set was randomly assigned. Each image was shown once, and all participants viewed the images in the same order. Each photo was 192 pixels in width and 256 pixels in height.

2.3 Procedure

All participants were first asked to fill out a questionnaire to determine previous exposure to Japanese animation. A mixed factorial design was then implemented. Participants were assigned to one of two adaptation groups (cartoon or live action) matched for gender and ethnicity. The cartoon group was assigned to watch the 50-min cartoon video, and the live action group was assigned to watch the 50-min live action video. Both groups rated test stimuli for attractiveness on a scale of 1 (unattractive) to 7 (attractive) before and after watching the show. The time interval between participants viewing the videos and rating the faces was about 1-2 min, during which time they received instructions.

The experiment was run on 24 inch iMacs. Programming was done with PHP and hosted on a protected server. Morphman was used to create test stimuli images. Microsoft Excel and SPSS 16.0 were used to organize and process the data collected.

3 Results

The data collected were analyzed by similar methods as in experiment 1 of Rhodes et al (2003). For each participant, ratings for all faces with the same incremental eye size were averaged together. Third-order polynomials were fitted to each participant's data and used to estimate the eye distortion level of the most attractive faces through finding the local maxima of the polynomial curves. Third-order polynomials were used

because second-order curves may be restrictive in shape and fourth or higher-order curves possess more complexity than necessary.

Data from participants were excluded if erratic patterns were present. In the cartoon group, there were twenty-two participants; nineteen participants' data were included with average $R^2 = 0.91$, and three were excluded. In the live-action control group, there were twenty-three participants; eighteen participants' data were included with average $R^2 = 0.83$, and five were excluded.

Overall, we found that participants exposed to cartoon stimuli preferred human faces with larger eyes after adaptation while participants exposed to the live action stimuli did not (figure 4). We conducted a two-way mixed analysis of variance (ANOVA) with adapting condition (cartoon or live action) as a between-participants factor and the mean distortion levels of the most attractive faces before and after adaptation as a repeated-measures factor. No significant main effects were found for the adapting condition ($F_{1,37} = 2.32$, p = 0.14), or the before and after condition ($F_{1,27} = 0.452$, p = 0.51). The interaction was significant ($F_{1,37} = 6.45$, p = 0.02). As predicted, the most attractive face shifted toward the adapting distortion of larger eyes after adaptation to cartoon faces. Gender and ethnicity were included as covariates and were not found to be significant. When excluded subject data where the distortion level for the most attractive face could be identified were included in the analysis, the interaction was still significant (p < 0.05). However, we felt that, in those excluded cases, this method of analysis was not appropriate because the polynomial fit was not suitable for the data points.

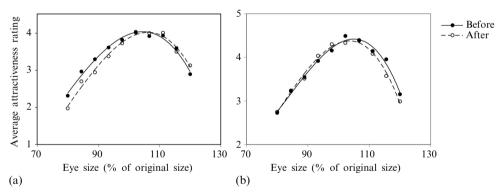


Figure 4. Mean attractiveness ratings as a function of eye size for the two groups before and after viewing videos. Ratings were made on a scale of 1 (unattractive) to 7 (attractive). There is a significant shift in preferences toward faces with larger eyes after adaptation to cartoons (a) but no significant difference in participants who viewed the live action show (b).

Planned comparisons showed no initial difference in preference for eye sizes between groups prior to video adaptation ($t_{35} = 0.44$, p = 0.67). In the test group that viewed cartoons, a paired *t*-test suggested that participants preferred faces with larger eyes after adaptation (M = 106.49, SD = 6.98) compared to before (M = 105.02, SD = 1.54), ($t_{18} = 2.12$, p = 0.048), showing a significant shift toward the adaptive distortion (figure 4a). This shift is seen for each of the 6 unique faces. In contrast, in the control group that viewed the live action show, there was no significant change in eye-size preference after adaptation (M = 101.63, SD = 5.10) compared with before (M = 104.13, SD = 5.64), ($t_{17} = -1.75$, p = 0.10), (figure 4b).

Also, all individuals preferred faces with larger than normal eyes, which corresponds with past findings suggesting that larger eyes are considered more attractive in female faces (Geldart et al 1999).

4 Discussion

This study demonstrates that adaptation to cartoon faces with large eyes is able to shift preferences for human faces toward larger eyes. It extends findings from past research on face adaptation in two ways. First, here videos have been used as adaptation stimuli rather than still pictures, offering a more naturalistic method of adaptation that corresponds more to real-life exposure. As a result, we are better able to generalize effects found from controlled laboratory settings to experience in everyday life.

Second, these results suggest that adaptation to cartoon faces can also affect preferences for human faces, expanding on previous findings that adaptation to distorted human faces affects preferences for novel human faces (Cooper and Maurer 2008; Peskin and Newell 2004; Rhodes et al 2003). This effect also points to a shared common face space between cartoon and human faces, building upon past work that suggests a common face space for all human faces (Jaquet and Rhodes 2008; Jaquet et al 2007a). Moreover, the use of videos reduces the likelihood that the effect resulted from primarily retinotopic adaptation as shown by Xu et al (2008). These findings support the idea of a general representation of faces.

Other research findings also support this conclusion. Electrophysiological cell recordings in macaque monkeys suggest that the middle face patch, which consists primarily of face-selective cells, is selective for human and cartoon faces, with response to cartoon faces about 83% of the response to human faces (Freiwald et al 2009). Likewise, imaging data in humans suggest that synthetic, cartoon-like faces also activate the fusiform face area, an area proposed to be specialized for face processing, about 84% as strongly as human faces (Loffler et al 2005).

In addition, this study may provide insight on why the uncanny valley exists. One possible explanation would be that human faces and other kinds of faces across the uncanny valley, like cartoon faces, are encoded by separate mechanisms. However, our results suggest that there are common mechanisms for processing different kinds of faces. Past work suggests that in addition to common mechanisms, there are dissociable mechanisms for face encoding (Jaquet and Rhodes 2008; Jaquet et al 2007a, 2007b; Little et al 2005, 2008; Rhodes et al 2003, 2004a, 2007; Webster et al 2004). Thus, the 'uncanny valley' phenomenon may actually stem from ambiguity in encoding between general and dissociable mechanisms. Because faces that are more fluently processed have been suggested to elicit more positive reactions (Reber et al 2004), faces and other representations that fall into the 'uncanny valley' may be less fluently processed, and as a result induce more negative reactions.

Finally, the findings from this study may have implications for media effects on preferences. Popular media have been shown to affect viewers' perceptions of beauty (Groesz et al 2002), likely from shifting perceptions using adaptation to specific ideals of beauty. Young people spend on average 16-17 hours weekly viewing television, beginning as early as the age of 2 years (Nielsen Media Research 1998), and cartoons are a key part of this media diet. That cartoons can change real-life preferences extends the scope of impact from various media sources.

Acknowledgments. We want to thank Deniz Kural for his assistance with task programming, and Shelley Carson for consultations throughout the project. This research was supported in part by a Harvard College Research Program grant to Haiwen Chen as well as NIH grants to Ken Nakayama (EY01362), Margaret Livingstone (EY16187), and Richard Russell (EY017245).

References

- Bernstein I H, Lin T D, McClellan P, 1982 "Cross- vs. within-racial judgments of attractiveness" Perception & Psychophysics 32 495-503
- Bronstad P, Russell R, 2007 "Beauty is in the eye of the beholder: Greater agreement on facial attractiveness among close relations" *Perception* **36** 1674–1681
- Cooper P A, Maurer D, 2008 "The influence of recent experience on perceptions of attractiveness" Perception 37 1216–1226

- Cunningham M, Roberts A R, Barbee A P, Druen P B, Wu C, 1995 "Their ideas of beauty are, on the whole, the same as ours': Consistency and variability in the cross-cultural perception of female physical attractiveness" *Journal of Personality and Social Psychology* **68** 261–279
- Dion K, Berscheid E, Walster E, 1972 "What is beautiful is good" Journal of Personality and Social Psychology 24 285-290
- Feingold A, 1992 "Good-looking people are not what we think" Psychological Bulletin 111 304-341
- Freiwald W, Tsao D, Livingstone M, 2009 "A face feature space in the macaque temporal lobe" Nature Neuroscience 12 1187–1196
- Geldart S, Maurer D, Carney K, 1999 "Effects of eye size on adults' aesthetic ratings of faces and 5-month-olds' looking times" *Perception* 28 361-374
- Groesz L M, Levine M P, Murnen S K, 2002 "The effect of experimental presentation of thin media images on body satisfaction: a meta-analytic review" *International Journal of Eating Disorders* **31** 1–16
- Jaquet E, Rhodes G, 2008 "Face aftereffects indicate dissociable, but not distinct, coding of male and female faces" *Journal of Experimental Psychology: Human Perception and Performance* **34** 101–112
- Jaquet E, Rhodes G, Hayward W, 2007a "Opposite aftereffects for Chinese and Caucasian faces are selective for social category information and not just physical face differences" *Quarterly Journal of Experimental Psychology* **60** 1457–1467
- Jaquet E, Rhodes G, Hayward W, 2007b "Race-contingent aftereffects suggest distinct perceptual norms for different race faces" *Visual Cognition* **16** 734-753
- Langlois J, Ritter J, Roggman L, Vaughn L, 1993 "Facial diversity and infant preferences for attractive faces" Annual Progress in Child Psychiatry and Child Development 1992: Selection of the Years Outstanding Contributions Eds M A Hertzig, E A Farber (Routledge, London) pp 27-41
- Langlois J H, Kalakanis L, Rubenstein A J, Larson A, Hallam M, Smoot M, 2000 "Maxims or myths of beauty? A meta-analytic and theoretical review" *Psychological Bulletin* **126** 390-423
- Leopold D A, O'Toole A J, Vetter T, Blanz V, 2001 "Prototype-referenced shape encoding revealed by high-level aftereffects" *Nature Neuroscience* **4** 89–94
- Leopold D A, Rhodes G, Muller K M, Jeffery L, 2005 "The dynamics of visual adaptation to faces" Proceedings of the Royal Society of London, Series B 272 897–904
- Little A, DeBruine L, Jones B, 2005 "Sex-contingent face after-effects suggest distinct neural populations code male and female faces" *Proceedings of the Royal Society of London, Series B* 272 2283
- Little A, DeBruine L, Jones B, Waitt C, 2008 "Category contingent aftereffects for faces of different races, ages and species" Cognition 106 1537-1547
- Loffler G, Yourganov G, Wilkinson F, Wilson H, 2005 "fMRI evidence for the neural representation of faces" *Nature Neuroscience* **8** 1386-1390
- MacDorman K, Green R, Ho C, Koch C, 2009 "Too real for comfort? Uncanny responses to computer generated faces" *Computers in Human Behavior* **25** 695–710
- Martin J, 1964 "Racial ethnocentrism and judgment of beauty" *Journal of Social Psychology* 63 59-63
- Mori M, 1970 "Bukimi no tani" Energy 7 33-35
- Nielsen Media Research, 1998 Report on Television (New York: Nielsen Media Research)
- Perrett D I, May K A, Yoshikawa S, 1994 "Facial shape and judgements of female attractiveness" *Nature* 368 239-242
- Peskin M, Newell F N, 2004 "Familiarity breeds attraction: effects of exposure on the attractiveness of typical and distinctive faces" *Perception* 33 147–157
- Reber R, Schwarz N, Winkielman P, 2004 "Processing fluency and aesthetic pleasure: is beauty in the perceiver's processing experience?" *Personality and Social Psychology Review* **8** 364–382
- Rhodes G, Jeffery L, Clifford C W, Leopold D A, 2007 "The timecourse of higher-level face aftereffects" Vision Research 47 2291-2296
- Rhodes G, Jeffery L, Watson T, Jaquet E, Winkler C, Clifford C, 2004a "Orientation-contingent face aftereffects and implications for face-coding mechanisms" *Current Biology* **14** 2119–2123
- Rhodes G, Jeffery L, Watson T L, Clifford C W, Nakayama K, 2003 "Fitting the mind to the world: face adaptation and attractiveness aftereffects" *Psychological Science* **14** 558-566
- Rhodes G, Jeffery L, Watson T L, Jaquet E, Winkler C, Clifford C W, 2004b "Orientationcontingent face aftereffects and implications for face-coding mechanisms" *Current Biology* 14 2119-2123
- Russell R, 2009 "A sex difference in facial pigmentation and its exaggeration by cosmetics" *Perception* **38** 1211 1219

- Samuels C A, Butterworth G, Roberts T, Graupner L, Hole G, 1994 "Facial aesthetics: babies prefer attractiveness to symmetry" *Perception* 23 823–831
- Seyama J, Nagayama R, 2007 "The uncanny valley: Effect of realism on the impression of artificial human faces" *Presence: Teleoperators and Virtual Environments* 16 337-351
- Slater A, Von der Schulenburg C, Brown E, Badenoch M, Butterworth G, Parsons S, Samuels C, 1998 "Newborn infants prefer attractive faces" *Infant Behavior and Development* **21** 345–354
- Steckenfinger S A, Ghazanfar A A, 2009 "Monkey visual behavior falls into the uncanny valley" Proceedings of the National Academy of Sciences of the USA 106 18362 – 18366
- Valentine T, 1991 "A unified account of the effects of distinctiveness, inversion, and race in face recognition" *Quarterly Journal of Experimental Psychology A* **43** 161–204
- Webster M A, Kaping D, Mizokami Y, Duhamel P, 2004 "Adaptation to natural facial categories" Nature 428 557-561
- Xu H, Dayan P, Lipkin R, Qian N, 2008 "Adaptation across the cortical hierarchy: Low-level curve adaptation affects high-level facial-expression judgments" *Journal of Neuroscience* 28 3374–3383

ISSN 1468-4233 (electronic)



www.perceptionweb.com

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.