The role of peripheral vision in the flashed face distortion effect

Jamie Bowden, David Whitaker, Matt J Dunn

School of Optometry and Vision Sciences, Cardiff University, Cardiff, UK

Correspondence to:

Dr Matt J Dunn

School of Optometry and Vision Sciences

Cardiff University

Maindy Road

Cardiff CF24 4HQ, Wales, UK

Tel: +44 (0)29 20870576

Email: DunnMJ1@cardiff.ac.uk

Keywords: peripheral vision, flashed face distortion effect

All data and code for stimuli (MATLAB) and analysis (R) are available at the UK Data Service's *ReShare* archive. See publication for download link.

Abstract

The flashed face distortion effect is a phenomenon whereby images of faces, presented at 4-5 Hz in the visual periphery, appear distorted. It has been hypothesised that the effect is driven by cortical, rather than retinal, components. Here, we investigated the role of peripheral viewing on the effect. Normally-sighted participants viewed the stimulus peripherally, centrally, and centrally with a blurring lens (to match visual acuity in the peripheral location). Participants rated the level of distortion using a visual analogue scale. Although optical defocus did have a significant effect on distortion ratings, peripheral viewing had a much greater effect, despite matched visual acuity. We suggest three potential mechanisms for this finding: increased positional uncertainty in the periphery, reduced deployment of attention to the visual periphery, and/or the visual crowding effect.

Introduction

The flashed face distortion effect (FFDE) is a perceptual phenomenon resulting from rapid presentation of eye- and mouth-aligned faces in the visual periphery. As the faces are presented sequentially, they appear increasingly distorted and deformed. The effect was first observed accidently by Tangen et al. in 2011, while scrolling through a set of eye-aligned Slovakian face images for an unrelated study (original stimulus available at https://youtu.be/wM6lGNhPujE). Upon further investigation, they noted that the effect was increased when the faces were viewed eccentrically and greatest in faces for which the dimensions of one or more of the facial features deviated significantly from the others in the set (e.g. if one face has a particularly large forehead, it appeared even larger and bulbous in shape). The effect was also reduced by rotating the faces by 180° (Tangen et al., 2011). Further investigation by Utz and Carbon (2015) showed that the effect decreases significantly if the faces used are from different ethnic origins or species. These findings suggest that a higher level of cortical, holistic perception is required for the FFDE to exist.

Despite the popularity of the effect after achieving second place in the Neural Correlate Society's *Best Illusion of the Year Contest* in 2012, little published literature exists on the effect. The mechanism by which the FFDE arises remains unclear.

Tangen et al. (2011) hypothesised that the FFDE is related to the face distortion aftereffect (FDAE), a phenomenon first discussed by Webster and MacLin (1999). The FDAE occurs when individuals are exposed to a distorted face image for an extended period of time. Then, when a 'normal' face is viewed immediately after, it appears distorted, with features in opposition to the initial, distorted, face. Further evidence for the possible similarity between the two effects was provided in a study by Wen and Kung (2014), who showed that similar visual areas of the brain are activated during fMRI experiments while viewing the FDAE and FFDE. However, the FDAE has been shown to be equally as strong in inverted and upright images of faces, suggesting it is a low-level effect mediated by cells in V1, as in the tilt aftereffect. On the other hand, subjects reported a much weaker FFDE when faces are inverted (Utz & Carbon, 2015). Furthermore, although it has been documented that there are greater aftereffects present in FDAE with increasing neural adaptation (Kovács, Cziraki, Vidnyánszky, Schweinberger, & Greenlee, 2008), Wen and Kung (2014) found that increasing adaptation led to a weaker FFDE response. The fMRI study concluded that in addition to early visual and face selective regions of the brain, "two additional groups, one for perceptual processing and two other subsystems relating to emotion and/or engagement" were also involved in the FFDE (Wen & Kung, 2014). This suggests that although the similar cortical locations process both effects, the neural adaptation that causes the FDAE (Kovács et al., 2008) is unlikely to cause the FFDE due to the brief stimulus exposure duration (200-250 ms per face). It is more likely that the FFDE is the result of top-down feedback from a higher perceptual processing area, as identified by Wen and Kung (2014).

Utz and Carbon (2015) suggest that the FFDE is based on configural processing, similar to typical face recognition. As the effect appears to rely on the exaggeration of the slight differences between each face (Tangen et al., 2011), it is most likely second-order configural information (i.e. information that cannot be isolated to a single feature, such as the spacing between the eyebrows and hairline) that is responsible for the effect, as this is the mechanism believed to be used for discriminating between faces (Le Grand, Mondloch, Maurer, & Brent, 2001; Young, Hellawell, & Hay, 2013). Piepers and Robbins (2012) discussed how the terms 'configural' and 'holistic' could often be interchanged when talking about the commonly accepted holistic/part-based model for face perception. McKone (2004) showed that the holistic part of this can be separated from part-based identification, as upright faces are still detectable in the visual periphery, despite the individual features being too blurred to recognise on their own, and furthermore, inverted faces are unrecognisable in the periphery due to reduced visual acuity (VA). This helps explain the significant decrease in the effectiveness of the FFDE when viewing inverted faces, as shown by Utz and Carbon

(2015), and adds further evidence that the effect is driven by a high-level perceptual mechanism, rather than a low-level neural or optical reason.

Here, we investigated the role of peripheral vision and visibility on the FFDE by comparing peripheral viewing to central viewing, and central viewing with optical defocus, in which the level of optical defocus produced a VA equivalent to that recorded under peripheral viewing. If the FFDE is similar both under peripheral viewing and with optical defocus, then the effect could be explained purely on the basis of reduced image quality. If on the other hand, the effect is greater under peripheral viewing, this would suggest an effect specific to peripheral vision; for example, an attentional or visual crowding mechanism, or an increase in positional uncertainty.

Method

Participants

Twelve healthy normally-sighted individuals were recruited on a voluntary basis from a University population, each with a best-corrected monocular VA of 0.00 logMAR or better and no known ocular pathology or abnormalities. The investigation was carried out in accordance with the Declaration of Helsinki; informed consent was obtained from the participants after explanation of the nature and possible consequences of the study. The Cardiff University School of Optometry and Vision Sciences Research Ethics Audit Committee granted approval for this study.

Materials

Stimuli were generated using the Psychophysics Toolbox extension for MATLAB (The MathWorks, Natick, MA, USA) (Brainard, 1997; Kleiner, Brainard, & Pelli, 2007; Pelli, 1997), and displayed on an iMac display of size 591 × 332 mm, running at a pixel resolution of 5120 × 2880 (Apple, Cupertino, CA, USA). All stimuli (FFDE and optotype VA task) were performed on the same screen at a distance of 1 m. The face images used for the FFDE were identical to those in the original study by Tangen et al. (2011), and were kindly provided by the author.

Procedure

The FFDE was initially demonstrated to participants with no input from the examiner, in order to familiarise them with the meaning of the term 'visual distortion' in the context of the study.

Participants had their monocular, foveal VA measured, to ensure they met the criteria of 'normal' vision, using single Sloan optotypes (Pelli, Robson, & Wilkins, 1988) with four crowding flankers of the same width as the optotype detail, spaced one bar width away (e.g. Figure 1).

Figure 1: Example optotype ('H') and flanking bars.

VA was determined according to Bailey and Lovie’s (1976) clinical protocol (five letters shown at each size in 0.10 logMAR steps; if four or more letters are incorrectly identified, the procedure ends; 0.02 is added to the result for each incorrect letter in the last set of five). All procedures were carried out in all participants with the right eye. The left eye was patched at all times. Peripheral VA was then measured at an eccentricity of 3° using the same test, but with the letters placed 3° to the right of a fixation target (Thaler, Schütz, Goodale, & Gegenfurtner, 2013). A spherical convex blurring lens was chosen to produce similar VA centrally (calculated to be within ±0.10 logMAR of the peripheral VA). Participants’ foveal VA was then re-measured but with the addition of this lens. If VA was not within ±0.10 logMAR of their peripheral VA, the lens was adjusted, and VA was re-measured; this process was repeated until an appropriate VA was recorded. Participants were not made aware of the aim of this procedure until the appropriate blurring lens was identified.

Participants then viewed the FFDE stimulus at a rate of 4 Hz (described as the optimum frequency to produce the effect by Utz and Carbon (2015), and also used by Wen and Kung [2014]) for 20 s. Faces were shown on a white background; only one face was shown at all times. The FFDE was viewed under three conditions. The ‘central’ condition relied simply on participants looking directly at the sequence of faces with no additional blurring lens. The ‘central blur’ condition again involved looking directly at the faces but with the addition of the blurring lens. The ‘peripheral’ condition involved viewing the faces peripherally (as in previous studies) at an eccentricity of 3°. Face images were 3.5° in width and 5° in height; therefore visual separation from the centre of the fixation cross to the edge of each face image was 1.75°. Participants were instructed to view a black 0.5° fixation target (Thaler et al., 2013), which was located 3° to the left of the centre of the face stimulus. A Latin square design was used to counterbalance stimulus presentation order across participants.

After each stimulus presentation, participants rated the level of distortion using pen and paper, on a Visual Analogue Scale (20 cm in length). A mark on the scale indicated the perceived level of distortion; the further to the right the cross was placed, the higher the level of distortion. The position of the recorded marks were physically measured with a ruler, and converted linearly into distortion values ranging from 0 (no distortion; far left of line) to 1 (far right).

Statistical analysis of the dataset was performed using the R Environment for Statistical Computing (R Core Team, 2012).

Results

Table 1 gives the VA and distortion ratings recorded for each participant under each viewing condition.

Table 1: VA and distortion ratings recorded for all participants under all conditions.

Participant	VA (logMAR)			Distortion rating		
	Central	Blur	Peripheral	Central	Blur	Peripheral
01	-0.10	0.48	0.48	0.32	0.60	0.74
02	-0.12	0.40	0.38	0.02	0.58	0.33
03	-0.30	0.40	0.50	0.05	0.29	0.39
04	-0.22	0.20	0.30	0.07	0.24	0.54
05	-0.02	0.48	0.50	0.04	0.44	0.82
06	0.08	0.40	0.50	0.01	0.22	0.08
07	-0.12	0.30	0.38	0.19	0.33	0.64
08	-0.02	0.38	0.38	0.16	0.05	0.76
09	-0.02	0.40	0.40	0.02	0.16	0.64
10	-0.12	0.50	0.50	0.02	0.11	0.82
11	0.00	0.58	0.48	0.08	0.29	0.49
12	-0.02	0.30	0.40	0.10	0.45	0.66

A boxplot summarising the ratings given under each viewing condition is presented in Figure 2.

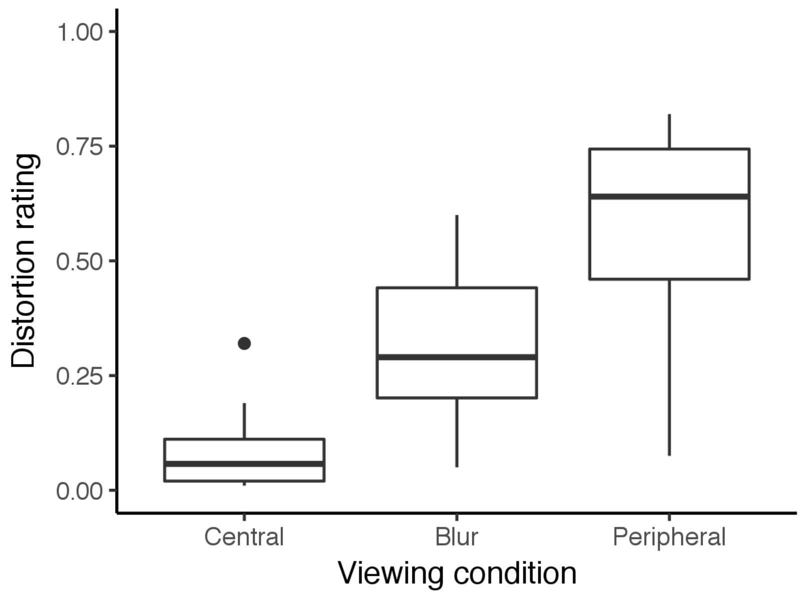

Figure 2: Boxplot showing distortion ratings for each viewing condition

A repeated-measures ANOVA was performed to determine the effect of viewing conditions on distortion ratings, which was found to be highly significant [$F(2,22) = 16.53$, $p < 0.001$]. A post-hoc Tukey test showed that distortion ratings differed significantly from one another in all viewing conditions (all Bonferroni-corrected $p \leq 0.03$). Peripheral viewing produced the highest distortion ratings (mean ± SD = 0.57 ± 0.22), followed by optical defocus (0.31 ± 0.17), followed by central viewing (0.09 ± 0.09). The mean difference between viewing with optical defocus vs. peripherally (0.26 difference) was significantly greater ($p = 0.03$) than the mean difference between optical defocus vs. central viewing (0.22 difference).

A Bayesian one-sample t-test was performed to determine whether central viewing of the FFDE (without defocus) was significantly different from zero. This resulted in a Bayes Factor (BF_{10}) of 8.05, indicating that even under central viewing, some distortion is perceived with the FFDE.

Discussion

This study aimed to investigate whether the FFDE can be explained simply by poor visibility, or whether other mechanisms are likely to cause the effect. Our findings indicate that peripheral vision is important in the effect, i.e. despite similar VA achieved by optical defocus, the effect is significantly weaker when viewed directly.

Despite the weaker effect, our results demonstrate that the FFDE – to some extent – is still elicited under central viewing, and that the effect can be increased significantly simply by introducing blur. A mean increase in distortion ratings of 0.22 was observed under blurred conditions, suggesting that optical defocus does have a small role to play in the effect. However, as peripheral viewing caused a further (significant) mean increase in distortion ratings of 0.26 (total increase of 0.48 as compared to central viewing), we argue that the effect is likely not solely driven by poor visibility, but that an effect specific to peripheral vision is also important. Below, we suggest three possible mechanisms for the effect of peripheral vision: positional uncertainty, reduced attention to the visual periphery, and the crowding effect.

Positional uncertainty

Due to increased spatial pooling by retinal ganglion cells in the visual periphery, signals received from these retinal regions have greater positional uncertainty (Hussain et al., 2015). This lack of spatial precision has the potential to allow for interpretations of images that are distorted to a greater extent than images presented at the fovea.

Despite the fact that participants were provided with an equivalent level of defocus under central viewing conditions to match their peripheral VA, even under conditions of optical defocus, *within* the blurred image, the central retina maintains the advantage of increased spatial precision. Therefore, it is possible that positional uncertainty *per se* could be responsible for the increase of the effect in peripheral vision.

Positional uncertainty in the visual periphery may also arise as a result of information compression. Rosenholtz, Huang and Ehinger (2012) describe how information in the periphery is compressed into a limited capacity channel. According to their model, information being recoded and decompressed in the brain could also result in spatial imprecision, providing greater potential for distortion/exaggeration of facial features when the information is interpreted.

Visual attention

Evidence suggesting a possible attentional mechanism can be found in the fMRI study conducted by Wen and Kung (2014), which concluded that amongst other brain regions, a potential subsystem relating to engagement could have been involved in the FFDE.

Although our paradigm explicitly required participants to attend to the visual periphery, covert visual attention is a highly unusual scenario in day-to-day life. It is feasible that even when attention is deployed to the visual periphery, the actual level of attention may not be high as when attending to centrally presented faces. Therefore, we speculate that reduced visual attention to peripherally-presented faces could explain our results, despite being explicitly instructed to attend solely to the faces.

The extent to which peripheral attention might mediate the FFDE could be investigated by observing the effect in the presence of visual and/or cognitive distractors. If lack of attention plays an important role in driving the FFDE, then one would expect distortion ratings to increase in the presence of distractors.

Visual crowding

Another quality that differs between central and peripheral vision is that of visual crowding, often thought of as an unavoidable combination of attributes of a target (such as shape or orientation) with those of other nearby objects. Visual crowding zones increase in extent at a much faster rate than VA with increasing eccentricity (Toet & Levi, 1992), hence peripheral viewing is particularly susceptible to crowding effects. Crowding is thought to be distinct from visual attention (Dakin, Bex, Cass, & Watt, 2009).

Facial recognition is susceptible to crowding of individual features (Martelli, Majaj, & Pelli, 2005; Pelli & Tillman, 2008); it is possible that a similar mechanism might be responsible for the increases in peripheral distortion observed in the present study. The shape of a mouth, for example, might be distorted by that of its neighbouring chin or nose, and the global result of all such effects might be the type of grotesque distortion experienced.

Visual crowding occurs not only for stimuli that are in close spatial proximity, but also for stimuli that are temporally close to one another. Yeshurun, Rashal and Tkacz-Domb (2015) investigated *temporal crowding* for letters by varying presentation rate. Observers most accurately identified letters with an interstimulus interval of 300 ms, corresponding to a presentation rate of 3.03 Hz (taking into account stimulus presentation duration). The temporal crowding effect was significantly increased at higher presentation rates (up to 7.69 Hz). This is not in line with the known effect of presentation rate on the FFDE, which has been shown to be optimal at 4 Hz (Utz & Carbon, 2015), suggesting that temporal crowding is not responsible for the FFDE. However, it is important to note that in the above study, stimuli were flashed briefly with a gap, rather than presented constantly.

The process of face identification often relies on recognising low contrast features. Contrast sensitivity for high and intermediate spatial frequencies diminishes in the visual periphery (Rovamo, Virsu, & Näsänen, 1978), potentially leading to certain facial features appearing or disappearing if they differ in contrast to a consecutively-presented face. This might contribute to increased FFDE distortions in the visual periphery. Future work could investigate the extent to which local contrast differences between consecutively-presented facial features affects distortion ratings.

One potential way to investigate the extent to which crowding is responsible would be to quantify the effect in amblyopic observers. A notable feature of amblyopic vision (aside from reduced central VA) is the presence of strong crowding effects at fixation. So not only is their central vision blurred (as in our experiment with induced defocus);

it also suffers from crowding effects (Kalpadakis-Smith, Tailor, Dahlmann-Noor, & Greenwood, 2016).

The fact that optical defocus alone caused an increase in distortion ratings may be due to difficulty recognising individual facial features under blurred conditions, switching instead to a more holistic face recognition process (McKone, 2004). This, in turn, could lead to any differences between the different faces appearing exaggerated and thus distorted.

The present study demonstrates that although optical defocus can mediate the FFDE, peripheral viewing is a more important factor in the level of facial distortions perceived. Increased positional uncertainty, a reduction in visual attention and/or increased visual crowding in the periphery may all help to explain the effect. Further work is necessary to clarify the role of these potential mechanisms.

Acknowledgements

The authors thank Jon Erichsen for loan of equipment used in this study, and Jeremy Guggenheim for statistical advice. We also thank two anonymous reviewers for their helpful comments and suggestions.

References

Bailey, I. L., & Lovie, J. E. (1976). New design principles for visual acuity letter charts. *American Journal of Optometry and Physiological Optics*, *53*(11), 740–5. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/998716

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176952

Dakin, S. C., Bex, P. J., Cass, J. R., & Watt, R. J. (2009). Dissociable effects of attention and crowding on orientation averaging. *Journal of Vision*, *9*(11), 28. https://doi.org/10.1167/9.11.28

Hussain, Z., Svensson, C.-M., Besle, J., Webb, B. S., Barrett, B. T., & McGraw, P. V. (2015). Estimation of cortical magnification from positional error in normally sighted and amblyopic subjects. *Journal of Vision*, *15*(2), 25–25. https://doi.org/10.1167/15.2.25

Kalpadakis-Smith, A., Tailor, V., Dahlmann-Noor, A., & Greenwood, J. (2016). The perceptual effects of crowding in amblyopic and peripheral vision. *Journal of Vision*, *16*(12), 237. https://doi.org/10.1167/16.12.237

Kleiner, M., Brainard, D., & Pelli, D. G. (2007). What’s new in Psychtoolbox-3?

Perception ECVP Abstract, *36*. https://doi.org/10.1068/v070821

Kovács, G., Cziraki, C., Vidnyánszky, Z., Schweinberger, S. R., & Greenlee, M. W. (2008). Position-specific and position-invariant face aftereffects reflect the adaptation of different cortical areas. *NeuroImage*, *43*(1), 156–164. https://doi.org/10.1016/j.neuroimage.2008.06.042

Le Grand, R., Mondloch, C. J., Maurer, D., & Brent, H. P. (2001). Early visual experience and face processing. *Nature*, *410*(6831), 890–890. https://doi.org/10.1038/35073749

Martelli, M., Majaj, N. J., & Pelli, D. G. (2005). Are faces processed like words? A diagnostic test for recognition by parts. *Journal of Vision*, *5*(1), 6. https://doi.org/10.1167/5.1.6

McKone, E. (2004). Isolating the special component of face recognition: peripheral identification and a Mooney face. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *30*(1), 181–197. https://doi.org/10.1037/0278-7393.30.1.181

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/9176953

Pelli, D. G., Robson, J. G., & Wilkins, A. J. (1988). The design of a new letter chart for measuring contrast sensitivity. *Clin Vision Sci*, *2*(3), 187–199.

Pelli, D. G., & Tillman, K. A. (2008). The uncrowded window of object recognition. *Nature Neuroscience*, *11*(10), 1129–1135. https://doi.org/10.1038/NN.2187

Piepers, D. W., & Robbins, R. A. (2012). A review and clarification of the terms "holistic", "configural", and "relational" in the face perception literature. *Frontiers in Psychology*, *3*, 559. https://doi.org/10.3389/fpsyg.2012.00559

R Core Team. (2012). R: a language and environment for statistical computing. Vienna, Austria. Retrieved from http://www.r-project.org/

Rosenholtz, R., Huang, J., & Ehinger, K. A. (2012). Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Frontiers in Psychology*, *3*, 13. https://doi.org/10.3389/fpsyg.2012.00013

Rovamo, J., Virsu, V., & Näsänen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature*, *271*(5640), 54–6. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/625324

Tangen, J. M., Murphy, S. C., & Thompson, M. B. (2011). Flashed face distortion effect:

grotesque faces from relative spaces. *Perception*, *40*(5), 628–630. https://doi.org/https://doi.org/10.1068/p6968

Thaler, L., Schütz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research*, *76*, 31–42. https://doi.org/10.1016/j.visres.2012.10.012

Toet, A., & Levi, D. M. (1992). The two-dimensional shape of spatial interaction zones in the parafovea. *Vision Research*, *32*(7), 1349–1357. https://doi.org/https://doi.org/10.1016/0042-6989(92)90227-A

Utz, S., & Carbon, C.-C. (2015). Is the Flashed Face Distortion Effect expertise-based? - a systematic experimental investigation. *Journal of Vision*, *15*(12), 147. https://doi.org/10.1167/15.12.147

Webster, M. A., & Maclin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review*, *6*(4), 647–653. https://doi.org/10.3758/BF03212974

Wen, T., & Kung, C.-C. (2014). Using functional magnetic resonance imaging to explore the flashed face distortion effect. *Journal of Vision*, *14*(12), 29.

https://doi.org/10.1167/14.12.29

Yeshurun, Y., Rashal, E., & Tkacz-Domb, S. (2015). Temporal crowding and its interplay with spatial crowding. *Journal of Vision*, *15*(3), 11. https://doi.org/10.1167/15.3.11

Young, A. W., Hellawell, D., & Hay, D. C. (2013). Configurational Information in Face Perception. *Perception*, *42*(11), 1166–1178. https://doi.org/10.1068/p160747n