RESPONSE TO KATZHENDLER AND WEINSHALL:

Initial visual degradation during development may be adaptive

Lukas Vogelsang[a], Sharon Gilad-Gutnick[b], Sidney Diamond[b], Albert Yonas[c], and Pawan Sinha[b,1]

We thank Katzhendler and Weinshall for their thought-provoking comment (1) on our paper (2). They argue that the computational simulations are insufficient to suggest that initially poor acuity may be an adaptive feature of visual development.

The logic of their argument is this: Our evaluation of DNN performance is based not on classification of high-resolution images exclusively, but a range of image resolutions. Katzhendler and Weinshall (1) argue that low-resolution images either are not ecologically relevant and should therefore be neglected or, if relevant, then are likely available at all developmental time points, enabling training with a mix of low- and high-resolution images. Using such a mix, they indicate, yields better performance than a regimen that progresses from low- to high-resolution images. We offer responses to both possibilities.

First, even setting aside their ecological significance, the ability to recognize low-resolution images serves as a valuable characterization of the visual system. As an analogy, consider our impressive ability to recognize severely compressed or contrast-chimeric faces (3, 4). These transformations are not found in the natural environment, and yet knowledge of the visual system's ability to tolerate them provides insights about the nature of internal representations that underwrite our recognition skills. By the same token, knowing that humans are proficient at low-resolution face recognition (5–8), we can infer that their underlying representational strategies must somehow be able to accommodate such image degradations. Any computational proposals that purport to explain human performance must, therefore, also account for good generalization across resolution levels. In this sense, our evaluation of classification performance across high and low resolutions is not an ad hoc choice, but an attempt to examine how well a given training regimen is able to explain the versatility of human observers.

Let us now consider the second possibility Katzhendler and Weinshall (1) bring up: If low-resolution images are ecologically important, they would be available early in development as well, allowing for a mixed-resolution training regimen. However, implementing a mixed-resolution training protocol, although straightforward computationally, may not be practicable in natural development. Endogenous factors that limit acuity (e.g., retinal immaturity) cannot be toggled to yield mixed-resolution experiences. The external environment may be unreliable and too variable across individuals, as a generator of these experiences. This is especially notable since the opportunities for experience-dependent change may be rather time limited (9). Environmental vagaries can be sidestepped by having a developmental progression that ensures consistent experience with low-resolution inputs. Furthermore, the image classification scenarios (e.g., motion-blur, haze, low-light conditions) that benefit from low-resolution experience may not themselves serve as good sources of training exemplars. Although it remains to be rigorously demonstrated, low-resolution images might serve as a common root to achieve high performance on a wide variety of potential transformations, while ignoring idiosyncrasies of those individual transformations.

To summarize, we believe that endogenous factors, which enforce initial experience with degraded inputs, may indeed serve an adaptive purpose. The low- to high-resolution progression may be a compromise between the goal of maximizing performance and the constraints of biological feasibility. We appreciate Katzhendler and Weinshall's (1) thoughtful letter and the opportunity to respond to their points.

1 G. Katzhendler, D. Weinshall, Potential upside of high initial visual acuity? *Proc. Natl. Acad. Sci. U.S.A.* **116**, 18765–18766 (2019).

2 L. Vogelsang *et al.*, Potential downside of high initial visual acuity. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 11333–11338 (2018).

3 S. Gilad-Gutnick, E. S. Harmatz, K. Tsourides, G. Yovel, P. Sinha, Recognizing facial slivers. *J. Cogn. Neurosci.* **30**, 951–962 (2018).

4 S. Gilad, M. Meng, P. Sinha, Role of ordinal contrast relationships in face encoding. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 5353–5358 (2009).

5 L. D. Harmon, The recognition of faces. *Sci. Am.* **229**, 71–82 (1973).

6 T. Bachmann, Identification of spatially quantized tachistoscopic images of faces: How many pixels does it take to carry identity? *Eur. J. Cogn. Psychol.* **3**, 85–103 (1991).

7 A. W. Yip, P. Sinha, Contribution of color to face recognition. *Perception* **31**, 995–1003 (2002).

8 P. Sinha, B. J. Balas, Y. Ostrovsky, R. Russell, Face recognition by humans: 19 results all computer vision researchers should know about. *Proc. IEEE* **94**, 1948–1962 (2006a).

9 N. Daw, *Visual Development* (Springer, 2014).