Social Vision: Breaking a philosophical impasse?

Josefa Toribio ICREA-UAB

Sensitivity vs. experience

Adams and Kveraga's general hypothesis is that our visual system has evolved to extract the meaning of social visual cues that convey information about other people's emotions and intentions. This swift visual understanding of social cues makes adaptive sense, they argue, since it prepares us for anticipating events and behaviours that are essential for oiling the wheels of our social life.

There is, however, an interesting tension in the findings that Adams and Kveraga review. Most of the evidence they collect shows top-down influences impacting functional integration of social cues at a very early stage of visual information processing, thus strengthening the idea of cognitive penetrability in social vision. Yet, the lower the level at which such top-down influences occur, the less relevant the evidence would seem to be for clarifying issues pertaining to the phenomenology of visual experience, since such low-level processing takes place outside conscious awareness.

We must thus distinguish two very different questions. The first is whether properties like being angry or being threatening are properties that the visual system is *sensitive* to. Most of the findings Adams and Kveraga (2015) review encourage an affirmative answer to this question. Especially interesting are the studies involving blindsight patients, whose response to emotional facial expressions presented in their unimpaired visual field was greatly speeded up when a stimulus consisting in a person's body displaying the same emotion was also presented in their impaired area. Confirming the role of shared social affordances in social vision—due to the low-level integration of visually dissimilar stimuli—these studies illustrate the existence of a feedforward integration mechanism that seems to process visual information about emotions even though the processing bypasses the primary visual cortex and hence conscious awareness.

The second, and more philosophically interesting question, is whether we can visually *experience* properties like being angry or being threatening—as opposed to seeming to visually experience them as a result of a cognitive event such as a perceptual judgment. This latter distinction fuels one of the most fraught debates in the philosophy of perception today. According to "rich" views about the contents of visual experience, we can indeed visually experience high-level properties such as being threatening, being angry or being a pine tree. These properties, it is claimed, are part of the sensory phenomenology of visual experiences. By contrast, according to "thin" or "sparse" views, visual experience can only involve low-level features, such as colour, shape, texture or movement.

Philosophical arguments on both sides of the debate have failed to settle this issue. Some have taken this to imply that there is no fact of the matter as to whether we should favour rich over sparse views, and suggest, but do not explore, the possibility that vision science could help resolve this philosophical impasse (Logue thanks Ophelia Deroy for this suggestion in her (2013)). In what follows, I would like to discuss Adams and Kveraga's work in social vision as a step in that direction. But first things first: the arguments.

Pretty much everyone agrees that we can visually experience low-level properties such as colour, shape, motion or size. Advocates of sparse views hold that these are the only properties we can visually experience. They do not deny that things can visually seem to us as having high-level properties. They deny that these seemings have any phenomenology of their own (e.g. Tye, 1995) or, any sensory, as opposed to cognitive, phenomenology (e.g. Lyons, 2005). According to a widespread version of the sparse view, things visually seem to us as having high-level properties due to perceptual judgments occurring downstream of visual consciousness.

Probably the most detailed argument in favour of rich views is Siegel's (2006, 2010) phenomenal contrast argument. Siegel starts off with the plausible assumption that there is a phenomenological difference between the overall experiences one has before and after acquiring a recognition capacity such as the capacity to identify pine trees. Let's call these overall experiences O1 and O2. If there is a phenomenological difference between O1 and O2, Siegel argues, then the most plausible explanation is

that the visual experiences, E1 and E2, which are parts of O1 and O2, differ in phenomenal character. She establishes this by discussing and eliminating two alternative explanations: that the altered phenomenology of O2 is due to the occurrence of cognitive state, e.g. a judgment, and that O1 and O2 differ in background phenomenology. She then argues that if E1 and E2 differ in phenomenal character, then the properties visually experienced while undergoing E1 and E2 are different. Since the low-level properties of the pine trees are the same in both E1 and E2, what you experience in E2, she concludes, is the property of being a pine tree. The same could be said for other high-level properties, like being a banana, being a table or being John Malkovich. Here, too, her argument is an argument to the best explanation, as she proceeds by discussing and rejecting two alternative accounts: there being some nonrepresentational feeling of familiarity in E2 and its representing gestalt properties as opposed to the high-level property of being a pine tree.

Perceived emotion

Can Adams and Kveraga's functional forecast model help to settle the debate between sparse and rich views? Let us focus on emotion. The model assumes two different pathways in the processing of visual emotional cues (Weisbuch & Adams 2012). The first prepares us to anticipate imminent physical danger and survival prospects in the environment through sensitivity to unattended emotional cues. When considering pre-attentive responses to e.g. emotion expressions on members of a different racial group, some studies show that white and black participants exhibit a negative affect when they are subliminally exposed to the joy of the members of the other race (Weisbuch & Ambady, 2008). In general, evidence involving this pathway suggests that we are visually sensitive to very basic, evolutionarily relevant, albeit still high-level, properties such as being eatable or being dangerous. The sensitivity to these properties, and not low-level ones, is easily explained in evolutionary terms. Quickly detecting the presence of danger, or food, or a possible mate makes evolutionary sense in a way that detecting the presence of shapes and colours does not. Inasmuch as sparse views assume the visual system to be sensitive only to low-level properties, these results could be used to help decide against them. But the interesting issue about phenomenology would still remain unresolved by data involving this pathway.

The second pathway assumed by the functional forecast model involves conscious attention and triggers specific expectations about the likely behaviour of particular individuals based, it is claimed, on the visual *perception* of their emotion expressions. Studies about the influence of emotion in gender recognition show, for instance, that androgynous faces with angry expressions are more likely to be perceived as male, while faces with expressions of joy or fear tend to be perceived as female. They also show that when subjects are exposed to both male and female faces expressing joy, anger, sadness, fear or neutral faces, female faces expressing anger took the longest to be identified (Hess et al. 2009). The suggestion is that the visual experience of emotion is what helps (or disrupts, as in the second study) the gender categorization task. Yet, evidence coming from categorical identification tasks is, by its very nature, extremely ill-fitted for establishing anything about the underlying mechanisms, and hence to determine whether emotional properties are part of the contents of the visual experience itself or our seeming to experience them is the result of a post-perceptual event.

Could it not be that what seems like visual perception of emotional properties is instead the result of a perceptual judgment that subjects make based on their visually experiencing the low-level properties of the facial display? That both the pre-attentive and the attention driven response pathways are shown to be effortless, automatic and unintentional processing routes would seem to speak initially against this interpretation. Yet, even though, paradigmatically, judgments are taken to be conscious events deliberately formed based on evidence or as a result of reasoning, the idea of a perceptual judgment seems to be far from this paradigm. Perceptual judgments often occur without us realizing it. They also tend to be effortless, automatic and unintentional, and so are other post-perceptual processes that have nothing to do with vision, such as semantic priming.

Data about speed of visual information processing or selective activation in different brain regions also fail to deliver the right sort of evidence. The first kind tends to involve subliminal exposure to stimuli, which again may speak in favour of sensitivity, but not phenomenology. Moreover, even if the activation in the relevant brain areas were due to unique processing of visual information, it seems difficult to determine, in many of these experiments, whether the selective activation is a

response to emotional properties as such or low-level visual properties like the squareness of the jaw, the shape of the face or the roundness of the eyes.

The tie-breaker factor for distinguishing between visual and post-perceptual processes like perceptual judgments seems to be that what we visually experience has a kind of irresistibility that comes from mental processes being "strongly and directly controlled by specific and subtle features of the visual input itself" (Scholl & Gao, forthcoming, p. 9). The most relevant results would thus be those that show subjects being inescapably caught by the specifically emotional nuances of the visual stimuli, even when they know that they are really irrelevant or that they could distract them from their explicit task. When the task is gender recognition, that it takes subjects much longer to identify angry female faces is, for instance, what would be expected if anger was modulating attention exogenously as part of the subjects' ongoing visual experience. This, together with the evolutionary and/or developmental salience of emotional properties found in the studies related to the non-attentional pathway, speaks in favour of there being this sort of irresistibility to the visual stimuli, which is the hallmark of visual processing. The data collected in support of Adams and Kveraga's functional forecast model thus suggest that we have compelling reasons to think that we can visually experience high-level properties, although perhaps just of the kind that matters for meeting our biological and social needs—data remain silent about properties like being John Malkovich. In particular, the findings suggest that the distinctive phenomenal character of experiences involving this sort of basic emotions and social traits is best explained by their being visually experienced as opposed to being brought about by perceptual judgments.

Funding

References

Adams, R. & Kveraga, K. (2015). Social vision: Functional forecasting and the integration of compound social cues. *Review of Philosophy and Psychology*. This number.

Hess, U. Adams, R.B., Jr., Grammer, K., & Kleck, R.E. (2009). Sex and emotion expression: Are angry women more like men? *Journal of Vision, 9,* 1-8.

Logue, H. (2013). Visual experience of natural kind properties: is there any fact of the matter? *Philosophical Studies*, 162: 1–12.

Lyons, J. (2005). Perceptual belief and nonexperiential looks. *Philosophical Perspectives*, 19: 237–256.

Siegel, S. (2006). Which properties are represented in perception? In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience*. Oxford: Oxford University Press.

Siegel, S. (2010). *The content of visual experience*. New York: OUP.

Scholl, B. J. & Gao, T. (forthcoming). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention*. MIT Press.

Tye, M. (1995). *Ten problems of Consciousness*. Cambridge, MA: MIT Press.

Weisbuch, M. & Adams, R. B. (2012). The functional forecast model of emotion expression. *Social and Personality Psychology Compass* 6/7: 499–514.

Weisbuch, M., & Ambady, A. (2008). Affective divergence: Automatic responses to others' emotions depend on group membership. *Journal of Personality and Social Psychology*, 95: 1063–1079.