

notion of content. This is ironic, since if Hutto and Myin are right, a properly construed enactivism has precisely the right shape to deal with representational content, namely by showing how to account for it where it is genuinely present. Besides, as “radical” as such enactivism might look, it should not be discredited by the fact that it breaks with prominent strands of research, or ways of talking in the sciences of cognition, including artificial intelligence, or if it runs counter to cherished intuitions. In fact, a leaf should be taken out of *The Embodied Mind*, where Varela, Thompson and Rosch (1991: 133) stated: “why do we assume that cognitive science cannot call into question these notions of representation and information not just philosophically but in its day-to-day research?”

Laura M. Nascimento is a PhD student at the Department of Philosophy of the State University of Campinas (UNICAMP), São Paulo, Brazil. Currently, she is a Visiting Research Student at the Centre for Philosophical Psychology, University of Antwerp, being funded by the São Paulo Research Foundation (FAPESP), BEPE #2016/20284-1.

Erik Myin is Professor of Philosophy at the University of Antwerp and director of the Centre for Philosophical Psychology. With Daniel Hutto he wrote *Radicalizing Enactivism: Basic Minds without Content* (2013), and *Evolving Enactivism: Basic Minds Meet Content* (2017). His research is supported by the Research Foundation Flanders (FWO), projects G048714N “Offline Cognition” and G0C7315N “Getting Real about Words and Numbers” and by the BOF Research Fund of the University of Antwerp, DOCPRO3 project “Perceiving affordances in natural, social and moral environments.”

RECEIVED: 19 OCTOBER 2017
ACCEPTED: 22 OCTOBER 2017

Active Vision: A Broader Comparative Perspective Is Needed

Lars Chittka

Wissenschaftskolleg zu Berlin,
Germany

lars.chittka/at/wiko-berlin.de

Peter Skorupski

St. George’s, University of London,
UK • pskorupski/at/sgul.ac.uk

> Upshot • We sympathize with the view that visual information is often acquired by active sampling of the environment, for example, through scanning movements. Not all vision is (en-)active, however – humans can capture important details of a visual scene at a glance, for example. The strategies of active sampling in various animals depend substantially on the structure of their visual systems and the representational capacities of their brains.

« 1 » We agree with Adrián Palacios, María-José Escobar and Esteban Céspedes that “the question about whether color is an objective feature of the world has lost relevance” (§35). However, we respectfully disagree that the enactivist perspective (Thompson, Palacios & Varela 1992) gave birth to this notion. As early as Hermann Grassmann’s (1853) mixture laws, it was clear that fundamentally different physical stimuli could generate the same perceptual states. Turning to comparisons with other animals, almost 150 years ago, John Strutt suggested:

“The assumed attractiveness of bright colors to insects would appear to involve the supposition that the color vision of insects is approximately the same as our own. Surely this is a good deal to take for granted.” (Strutt 1874: 6)

This fundamentally anticipates the idea that animals use spectral information in different ways, depending on ecology and sensory apparatus, as experimentally confirmed first by John Lubbock’s (1882) discovery of UV sensitivity in insects. There is much

more variety of color vision systems across animals than Palacios, Escobar & Céspedes acknowledge: even within the vertebrates, numbers of spectral receptor types range from one to five, with some species of invertebrates having substantially greater diversity still (Skorupski & Chittka 2011; Thoen et al. 2014). There are at least as many “colored worlds” as there are spectral receptor systems across animals.

« 2 » Notions of links between action and perception likewise date back to the 19th century; see, e.g., the explorations by Johannes Purkinje (1825) to explain why active, voluntary movement of the eye does not cause a perception of movement of the visual scene, while pressing the eye with a finger does induce perceived motion. This phenomenon of the “efference copy” (von Holst & Mittelstaedt 1950) has been studied neurophysiologically in crustaceans (Skorupski & Sillar 1986) and insects, where its neural underpinnings are now understood in unparalleled detail (Kim, Fitzgerald & Maimon 2015). Spontaneous turning movements in tethered flies could be conceived of as active visual probing of the environment, and the perception-as-action perspective fits comfortably with studies on spontaneous actions and decision-making with small brains (Brembs 2011). Whether this sweeps away the theoretical need for representations in understanding cognition remains perhaps questionable; the contextualist notion of representation in a weak sense, put forward by the authors of the target article, strikes us as useful.

« 3 » Humans are capable of identifying certain elements of visual scenes at a glance (Thorpe, Fize & Marlot 1996; Kirchner & Thorpe 2006), using parallel search (Wolfe 2000). The rapid categorization of the kind seen in humans and monkeys might rely on pre-attentive feed-forward processing of low-level features (Joubert et al. 2009; Walker, Stafford & Davis 2008). Extraction of image components at a glance in primates implies a time course that seems compatible with only a single sweep of sensory information along the known delays of the visual system (VanRullen & Thorpe 2002). This indicates that an active exploration of the visual environment (e.g., by saccades or active head or eye movements) is at least not necessary for all visual processing.

« 4 » In some animals a link between action and perception has been demonstrated empirically. For example, pigeons acquire stereotyped approach routes when learning visual pattern discrimination tasks, and fail at these tasks when they are prevented from using their accustomed route (Dawkins & Woodington 2000). In other behavioral contexts, pigeons actually move their heads in characteristic ways (“head-bobbing”) to stabilize their image during forward locomotion (Theunissen & Troje 2017). Stomatopod crustaceans are an extreme case since their many types of spectral photoreceptors are arranged in tiers across their eyes – hence their only possibility to obtain color information is by making scanning movements, pointing the different spectral receptor types at the target sequentially (spectral scanning; Thoen et al. 2014). Active vision has also been suggested to play a role in color perception of bees (Skorupski & Chittka 2011). As ommatidia contain a variety of different sets of color receptor types (Spaethe & Briscoe 2005; Wakakuwa et al. 2005), a single-colored object might be perceived as consisting of multiple pixels each with different colors – unless insects move their eyes over the object to generate a temporal integration (Skorupski & Chittka 2011). This could explain why bees, while they can discriminate distinct colors that are only briefly flashed (25ms), need longer integration times (50ms) for reliable discrimination when colors are similar (Nityananda, Skorupski & Chittka 2014).

« 5 » It could be that a single “sensory snapshot” is less useful for a miniature brain simply because there is less capacity for off-line processing (processing involving significant neural delays). In active perception, the time taken to build up a picture of the environment will depend on how much can be sampled at once, and in a miniature brain there may be less capacity to process a single sensory snapshot, and consequently a greater dependence on continuous online sampling of the scene. It is possible that reduced representational capacity (in the weak sense of representation intended by Palacios, Escobar & Céspedes) may lead to fundamentally different perceptual sampling strategies in large and small brains. However, in certain highly active miniature-brained animals, such as

flies and bees, reduced representational capacity may be compensated for by higher real-time sampling of their environment (Chittka & Skorupski 2011).

« 6 » In small nervous systems, active vision could compensate for limitations imposed by limited detail provided by reduced receptor arrays. With a dynamical-systems approach it is possible to show that an array of only seven photoreceptors can classify circles v squares, given self-generated movements (Beer 2003). In the biological realm, *Drosophila* larvae, with only 12 photoreceptors, can distinguish movement of conspecifics from other visual signals (Justice et al. 2012). Moving to bees, we find eyes with image-forming optics, but the 8,000 ommatidia of a bee’s eye would seem to provide rather a poor image (compared to that generated by millions of rods and cones in the human eye). On the other hand, the visual learning abilities discovered in bees seem rather remarkable. Could it be that active vision strategies enable bees to overcome the poor spatial resolution implied by the optics of the compound eye? Indeed, there is now evidence that bees require active scanning movements for all but the most basic visual discriminations (Nityananda, Skorupski & Chittka 2014).

« 7 » In an animal with limited storage or representational capacity, part of image representation could be moved to its environment, using active vision. For example, a discrimination task of square versus circle could be conventionally solved by storing images of the two shapes (either a pixel-by-pixel image or some compressed representation that highlights key features) and comparing them. Alternatively, to ease demands in representational capacity, the task could be solved by comparing what the shapes induce you to do. For example, if you have a habit of scanning contrasting boundaries during approach flight (which bees do), then it would suffice to scan only a tiny section of the shape to make the discrimination. The circle then becomes “that which makes me move in an arc” and the square “that which makes me move in a straight line.”

« 8 » In conclusion, we agree that in many animals and many scenarios, to perceive something is to actively sample it, and that perception is constituted in implicit

knowledge (accumulated both phylogenetically and ontogenetically) of how the properties of the sensory motor apparatus result in changes in the sensory input during perceptual sampling (for example, eye movements in humans). This account places more of the burden of perception in the world itself, rather than in putative mental representations. However, the extent to which this applies depends on the task, as well as the particular visual system and the representational capacity of the animal; large-brained animals such as humans might in some cases bypass the need for active sampling, whereas small-brained animals such as bees are strictly dependent on it.

Lars Chittka is a recipient of the Royal Society Wolfson Research Merit Award, an Advanced Fellowship from the European Research Council and the Lesley Goodman Award of the Royal Entomological Society. He is an elected Fellow of the Linnean Society, the Royal Entomological Society and the Royal Society of Biology. Chittka is an Editor of *PLoS Biology*, a member of the Faculty of 1000, and was a Panel Chairman for the European Research Council (2010–2013). He is the founder of the Psychology Department at Queen Mary University of London, where he is a Professor of Sensory and Behavioural Ecology.

Peter Skorupski is a neurobiologist with a long-standing interest in sensorimotor integration; in particular, central modulation of sensory information by motor context. More recently his research interests have focused on comparative aspects of color vision and photoreceptor response dynamics. Skorupski lectures in the Institute of Medical and Biomedical Education, St George’s, University of London.

RECEIVED: 20 OCTOBER 2017

ACCEPTED: 26 OCTOBER 2017