Grouping and the Metaphysics of High-Level Perceptual Experience

High-levelists hold that perceptual experiences of high-level properties (such as natural or artifactual kinds) have their own phenomenal character (Bayne, 2009; Burnston, 2023; Cavedon-Taylor, 2021; Fish, 2013; Nanay, 2011; Ransom, 2020; Siegel, 2011; Stokes, 2018). There is, on this view, something it is like to visually experience a tulip, over and above the experience of its shape, color, texture, and other low-level properties.

Presently, there are two broad approaches to the metaphysics of perceptual experience in general. According to the 'what'-approach, phenomenal character is fully constituted by *what* objects and properties are experienced. Paradigm examples include strong versions of naïve realism (Campbell, 2006; Fish, 2009; Sethi, 2025), and of representationalism (Dretske, 1995; Pautz, 2021; Tye, 1995). According to the 'how'-approach, phenomenal character is constituted, at least in part, by *how* objects and properties are experienced, e.g., by ways of perceiving (French & Phillips, 2020) or appearing (Beck, 2019), by mental paint (Block, 1996; Papineau, 2021), or by (affective) attitudes (de Vignemont, 2023; Jacobson, 2021).

In this talk, based on evidence about grouping in Multiple Object Tracking (MOT), we argue that high-levelism conflicts with the what-approach, but not with the how-approach. Thus, a viable form of high-levelism must adopt a how-approach.

In MOT paradigms, participants are asked to track a subset of moving targets among distractors. Tracking performance improves when targets share a low-level property that distractors lack, such as a common color or shape (Feria, 2012; Störmer et al., 2011). These improvements can be interpreted as the result of perceptual grouping (cf. Yantis, 1992): targets are grouped by subjects to form a unique dynamic figure which is easier to track than separated targets. Such grouping is facilitated when targets are similar to each other and dissimilar from nontargets. Crucially, Wei et al. (2016) show that similar improvements can occur when targets share a high-level property, suggesting that high-level perception can also support grouping.

However, in MOT, grouping based on high-level properties differs strikingly from grouping based on low-level properties. The latter is automatic and irresistible: it occurs even when it is task-irrelevant and disrupts performance, by competing with the correct grouping required for the task (Erlikhman et al., 2013; Wang et al., 2016). By contrast, the former is resistible and under voluntary control: when a high-level property is task-irrelevant—such as when some targets and some distractors have the same high-level property—participants can deliberately avoid high-level grouping (Wei et. al., 2018). This behavioral result is corroborated by neuroimaging evidence that high-level grouping recruits brain regions associated with top-down attentional control (Wei et. al., 2017).

This result has important phenomenological consequences. As a first rough pass (to be refined in the talk), perceptual grouping is driven by experienced object similarity, in accordance with the Gestalt law of similarity: all else being equal, objects that are experienced

as similar are automatically grouped together (Wagemans et al., 2012; Wertheimer, 1923). Thus, since perceiving a high-level property shared by several objects does *not* automatically lead to grouping of these objects, it follows that there is no experience of high-level similarity between the objects in question. Assuming high-levelism, i.e., that the experiences in question have similar high-level phenomenal character, we get the following premise:

> Premise 1: High-level phenomenal similarity between experiences does not imply an experience of high-level similarity between the objects of these experiences.

We add the following premise, which we argue is highly plausible upon reflection:

> Premise 2: If high-level phenomenal character is fully constituted by what is experienced (i.e., the what-approach is true), then high-level phenomenal similarity between experiences implies an experience of high-level similarity between the objects of these experiences.

The two premises together entail that the what-approach is false.

Now, suppose that high-level phenomenal character is constituted by mental paint, attitudes, or ways of perceiving, in accordance with the how-approach. As a result, high-level phenomenal similarity between experiences need not imply an experience of high-level similarity between the objects of these experiences. Thus, the how-approach is compatible with Premise 1 and therefore not threatened by our argument.

In support of this compatibility claim, we develop the following analogy with attitudinal theories of affective perception (de Vignemont, 2023; Jacobson, 2021). Molly fears spiders and snakes. Her perceptual experiences of spiders and snakes are similar with respect to fear. On the attitudinal approach, the phenomenal similarity between her experiences lies entirely in the affective attitudes she has towards the two animals, not in any experienced property of the animals. Consequently, the phenomenal similarity in question need not imply that she experiences fear-related similarity between the animals. We argue that the how-approach to high-level phenomenology yields the same result: high-level phenomenal similarity (between experiences) need not involve an experienced high-level similarity between the perceived objects themselves.

We next deal with an important objection. When arguing for Premise 1, we assumed that the lack of automatic grouping in the high-level case implies a lack of experience of high-level similarity between objects. It might be objected that high-level grouping fails to occur automatically not because subjects do not experience high-level similarity between objects, but because the grouping mechanism itself is low-level and consequently insensitive to such similarities.

Our response challenges the intelligibility of this objection. Consider a subject who experiences a set of blue circles surrounded by red circles, and who consequently experiences the blue circles as similar to each other (and as dissimilar from the red circles). In such a case, the subject cannot help but experience the figure constituted by the blue items – which implies their grouping. To suppose that all this experiential similarity (and dissimilarity) is present, yet the figure is not experienced at all, is—we submit—hard to make sense of. On this basis, we argue that the objection mistakenly treats grouping as an optional add-on to experienced similarity between objects.

References

Bayne, T. (2009). Perception and the Reach of Phenomenal Content. *The Philosophical Quarterly (1950-)*, *59*(236), 385–404.

Beck, O. (2019). Rethinking Naive Realism. *Philosophical Studies*, *176*(3), 607–633. https://doi.org/10.1007/s11098-018-1030-x

Block, N. (1996). Mental Paint and Mental Latex. *Philosophical Issues*, *7*, 19–49. https://doi.org/10.2307/1522889

Burnston, D. C. (2023). How to Think About Higher-Level Perceptual Contents. *Mind and Language*, *38*(5), 1166–1186. https://doi.org/10.1111/mila.12446

Campbell, J. (2006). *Reference and Consciousness*. Clarendon Press.

Cavedon-Taylor, D. (2021). High-Level Perception and Multimodal Perception. In H. Logue & L. Richardson (Eds.), *Purpose and Procedure in Philosophy of Perception* (p. 0). Oxford University Press. https://doi.org/10.1093/oso/9780198853534.003.0008

de Vignemont, F. (2023). Fifty Shades of Affective Colouring of Perception. *Australasian Journal of Philosophy*, *101*(1), 1–15. https://doi.org/10.1080/00048402.2021.1965176

Dretske, F. (1995). *Naturalizing the mind*. MIT Press.

Erlikhman, G., Keane, B. P., Mettler, E., Horowitz, T. S., & Kellman, P. J. (2013). Automatic feature-based grouping during multiple object tracking. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(6), 1625–1637. https://doi.org/10.1037/a0031750

Feria, C. S. (2012). The effects of distractors in multiple object tracking are modulated by the similarity of distractor and target features. *Perception*, *41*(3), 287–304. https://doi.org/10.1068/p7053

Fish, W. (2009). *Perception, Hallucination, and Illusion*. Oxford University Press.

Fish, W. (2013). High-level properties and visual experience. *Philosophical Studies*, *162*(1), 43–55.

French, C., & Phillips, I. (2020). Austerity and Illusion. *Philosophers' Imprint*, *20*(15), 1–19.

Jacobson, H. (2021). The Role of Valence in Perception: An ARTistic Treatment. *The Philosophical Review*, *130*(4), 481–531. https://doi.org/10.1215/00318108-9263939

Nanay, B. (2011). Do We See Apples as Edible? *Pacific Philosophical Quarterly*, *92*(3), 305–322. https://doi.org/10.1111/j.1468-0114.2011.01398.x

Papineau, D. (2021). *The metaphysics of sensory experience*. Oxford University Press.

Pautz, A. (2021) *Perception*. Routledge.

Ransom, M. (2020). Expert Knowledge by Perception. *Philosophy*, *95*(3), 309–335. https://doi.org/10.1017/S0031819120000157

Sethi, U. (2025). A (Qualified) Defense of Diaphaneity. In *The Relational View of Perception*. Routledge.

Siegel, S. (2011). *The Contents of Visual Experience*. Oxford University Press. http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195305296.001.0001/acprof-9780195305296

Stokes, D. (2018). Rich Perceptual Content and Aesthetic Properties. In A. Bergqvist & R. Cowan (Eds.), *Evaluative Perception*. Oxford University Press.

Störmer, V. S., Li, S.-C., Heekeren, H. R., & Lindenberger, U. (2011). Feature-based interference from unattended visual field during attentional tracking in younger and older adults. *Journal of Vision*, *11*(2), 1. https://doi.org/10.1167/11.2.1

Tye, M. (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. MIT Press.

Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., & von der Heydt, R. (2012). A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychological Bulletin*, *138*(6), 1172–1217. https://doi.org/10.1037/a0029333

Wang, C., Zhang, X., Li, Y., & Lyu, C. (2016). Additivity of Feature-Based and Symmetry-Based Grouping Effects in Multiple Object Tracking. *Frontiers in Psychology*, *7*, 657. https://doi.org/10.3389/fpsyg.2016.00657

Wei, L., Zhang, X., Li, Z., & Liu, J. (2018). The semantic category-based grouping in the Multiple Identity Tracking task. *Attention, Perception, & Psychophysics*, *80*(1), 118–133. https://doi.org/10.3758/s13414-017-1420-8

Wei, L., Zhang, X., Lyu, C., Hu, S., & Li, Z. (2017). Brain activation of semantic category-based grouping in multiple identity tracking task. *PLOS ONE*, *12*(5), e0177709. https://doi.org/10.1371/journal.pone.0177709

Wei, L., Zhang, X., Lyu, C., & Li, Z. (2016). The Categorical Distinction Between Targets and Distractors Facilitates Tracking in Multiple Identity Tracking Task. *Frontiers in Psychology*, *7*. https://doi.org/10.3389/fpsyg.2016.00589

Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psycologische Forschung*, (4), 301–350.

Yantis, S. (1992). Multielement visual tracking: Attention and perceptual organization. *Cognitive Psychology*, *24*(3), 295–340. https://doi.org/10.1016/0010-0285(92)90010-Y