Tweaking the concepts of perception and cognition

doi:10.1017/S0140525X15002733, e232

Ned Block
Department of Philosophy, New York University, New York NY 10003.
ned.block@nyu.edu
http://www.nyu.edu/gsas/dept/philo/faculty/block/

Abstract: One approach to the issue of a joint in nature between perception and cognition is to investigate whether the concepts of perception and cognition can be tweaked to avoid direct, content-specific effects of cognition on perception.

Firestone & Scholl (F&S) give lip service to the claim that the real issue is that of a joint in nature between cognition and perception, but their focus in practice is on denying "true" top-down effects. For this reason they neglect the need to tweak the concepts of perception and cognition in order to home in on the joint—if there is a joint. I suggest a reorientation toward the question of whether direct, content-specific effects of cognition on perception can be eliminated by such tweaking.

I will explore this idea with regard to what is perhaps the most dramatic of the known top-down effects of cognition on perception. It has been shown that mental images can be superimposed on percepts, creating a combined quasi-perceptual state. Brockmole et al. (2002) used a "locate the missing dot" task in which the subject's task was to move a cursor to a missing dot in a 5-by-5 array. If a partial grid of 12 dots appeared briefly followed in the same place in less than 50 ms by another partial grid of 12 different dots (that stays on the screen until the response), subjects were able to fuse the two partial grids and move the cursor to the missing dot, remembering nearly 100% of the dots on the first array. (If the subject erroneously clicked on a space in which there was a dot in the first array, that was counted as forgetting a dot in the first array.) However, if the second array was delayed to 100 ms, subjects' ability to remember the first array fell precipitously (from nearly 12 dots down to 4 dots).

Brockmole et al. explored extended delays – up to 5 seconds before the second array appeared. The amazing result they found was that if the second array of 12 dots came more than 1.5 seconds after the first array had disappeared, subjects became very accurate on the remembered dots. (See Fig. 1, following.) Instructions encouraged them to create a mental image of the first array and superimpose it on the array that remained

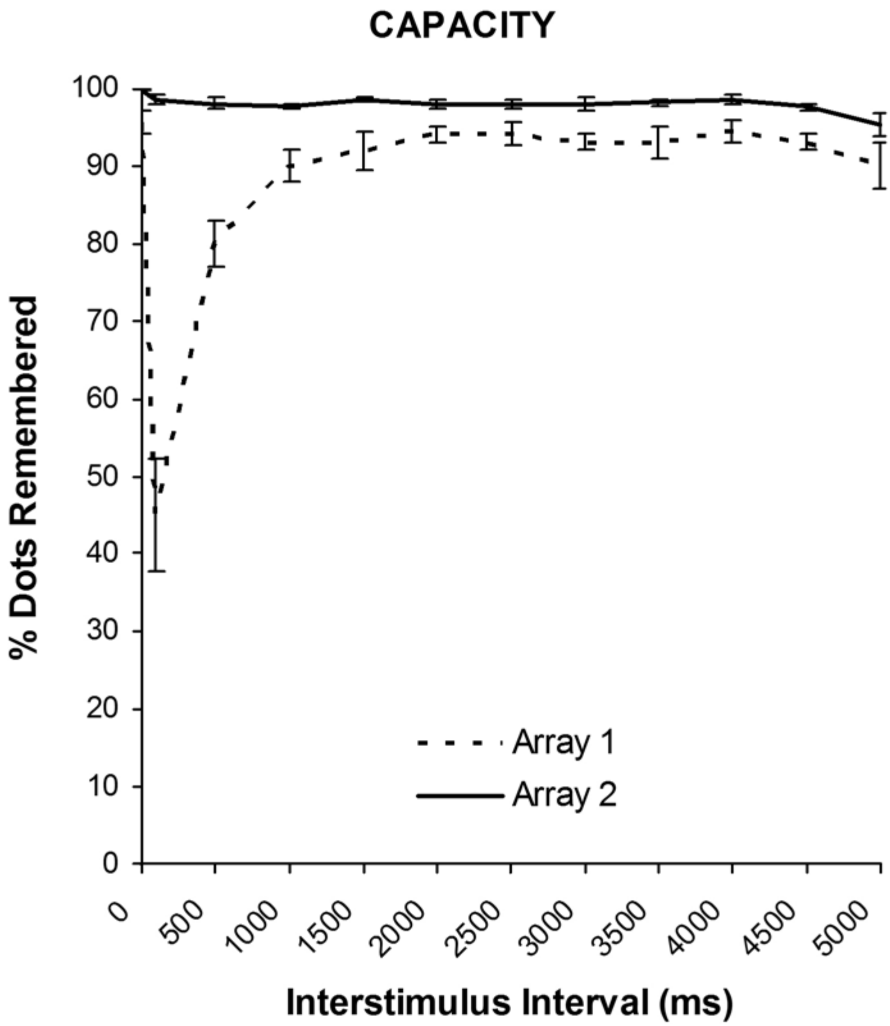

Figure 1 (Block). A 12-dot partial grid appears, then disappears. Another 12-dot partial grid appears – in the same place – up to 5 seconds later (i.e., after an interstimulus interval of up to 5 seconds). The second grid stays on the screen until the subject responds. The subject's task is to form a mental image of the first grid, superimposing it on the second grid so as to identify the dot missing from both grids. The dark line shows that subjects rarely click on an area that contains a dot on the screen. The dashed line shows that after a 1.5 second interval between the grids, subjects also rarely click on an area in which there was a dot in the remembered array. Thanks to James Brockmole for redrawing this figure.

on the screen. Brockmole et al. analyzed the subjects' errors, finding that subjects remembered more than 90% of the dots that they saw for up to 5 seconds. Further, the 1–1.5 seconds required to generate the mental image fits with independent estimates of the time required to generate a mental image (Kosslyn et al. 2006). Kosslyn and colleagues replicated this result with different methods (Lewis et al. 2011).

Those results demonstrate a direct content-specific effect of cognition on perception. And effects of this sort are involved in many of the phenomena used to criticize the idea of a joint in nature between cognition and perception. Many of the effects of language and expectation on perception that opponents of a joint in nature (e.g., Lupyan 2015a) cite seem very likely to involve superimposition of imagery on perception. For example, Lupyan notes that in a visual task of identifying the direction of moving dots, performance suffers when the subject hears verbs that suggest the opposite direction. Hearing a story about motion can cause motion aftereffects (Dils & Boroditsky 2010b), a point that enhances the plausibility that the result Lupyan notes involves some sort of combination of imagery and perception.

Another example: Delk and Fillenbaum (1965) showed that when subjects are presented with a heart shape and asked to adjust a background to match it, the background they choose is redder than if the shape is a circle or a square. As Macpherson (2012) points out, there is evidence that subjects are forming a mental image of a red heart and superimposing it on the cutout. Macpherson regards that as a case of "cognitive penetration." But if this is cognitive penetration, why should we care about cognitive penetration? Mental imagery can be accommodated to a joint between cognition and perception by excluding these quasi-perceptual states, or alternatively, given that imagery is so slow, by excluding slow and effortful quasi-perceptual states.

In expert perception – for example, wine-tasting, or a doctor's expertise in finding tumors in x-rays – we find another illustration of the benefit of using direct, content-specific, top-down effects of cognition on perception to tweak the category of perception. Opponents of a joint (Lupyan 2015a; Vetter & Newen 2014) often cite expert perception. But, as Fodor noted, the category of perception need not be taken to include perceptual learning (1983). The concept of perception can be shrunk so as to allow diachronic cognitive penetration in perceptual learning while excluding synchronic cognitive penetration.

Moving to a new subject: F&S's treatment of attention is inadequate. Oddly, they focus on spatial ("peripheral") attention and treat attention as akin to changing the input. However, Lupyan (2015a) is right that feature-based attention operates at all levels of the visual hierarchy (e.g., both orientations and faces) and so is not a matter of operation on input. Taking direction of motion as an example, here is how feature-based attention works: There is increased gain in neurons that prefer the attended direction; there is suppression in neurons that prefer other directions; and there is a shift in tuning curves toward the preferred direction (Carrasco 2011; Sneve et al. 2015). Although feature-based attention is directed by cognition, it may in many cases be mediated by priming by a sample of the feature to be attended (Theeuwes 2013). These effects cannot be eliminated by tweaking the concepts. But contrary to Lupyan (2015a), feature-based attention works by well-understood *perceptual mechanisms*, and there is not the slightest evidence that it directly modulates conceptual cognitive representations. This is a direct content specific effect of cognition on perception but what we know of the mechanisms by which it operates gives no comfort to opponents of a joint in nature between cognition and perception.

Finally, although I have been critical of F&S, I applaud their article for its wonderful exposé of the pitfalls in experimental approaches to the joint in nature between cognition and perception.

Acting is perceiving!

doi:10.1017/S0140525X15002460, e233

Rouwen Cañal-Bruland,[a] John van der Kamp,[b] and Rob Gray[c]

[a] *Institute of Sport Science, Department of Sport Psychology, Friedrich-Schiller-University Jena, 07749 Jena, Germany;* [b] *Department of Human Movement Sciences, Vrije Universiteit Amsterdam, 1081 BT Amsterdam, The Netherlands;* [c] *Human Systems Engineering, Arizona State University–Polytechnic, Mesa, AZ 85212.*

rouwen.canal.bruland@uni-jena.de **j.vander.kamp@vu.nl**
robgray@asu.edu
www.spowi.uni-jena.de/en/Sportpsychologie.html
https://vu-nl.academia.edu/johnvanderkamp
https://www.researchgate.net/profile/Rob_Gray/

Abstract: We challenge Firestone & Scholl's (F&S's) narrow conceptualization of what perception is and – most important – what it is for. Perception guides our (inter)actions with the environment, with attention ensuring that the actor is attuned to information relevant for action. We dispute F&S's misconceived (and counterfactual) view of perception as a module that functions independently from cognition, attention, and action.

Consider the following: An MLB pitcher throws a ball toward you that you need to hit. Presuming that most of our readers count as novices or, at best, less-skilled batters, then ample evidence indicates that – even if the pitcher performed the throw with the exact same movements, and hence would provide the exact same kinematic information as when he faced a professional pinch-hitter – you most likely would not (be able to) attune to and use the same information as the expert (Mann et al. 2007). The events with which you interact, the pitcher's throwing motion and the ball's flight, are identical and thus induce the very same patterns in the optic array, yet the information you attune to for guiding your batting action would be crucially different from the information the expert attunes to and uses. To paraphrase F&S, this is *the* perfect example for no change in visual input but a dramatic change in visual perception. It underlines our point that perception can be understood only by considering its primary purpose: to guide our actions in and interactions with the environment (Gibson 1979/1986; see also Cañal-Bruland & van der Kamp 2015).

Sticking with the batting example, there is no doubt that you and the professional pinch-hitter differ with respect to movement skills or action capabilities. Consequently, if the ball were to approach you and the expert player at the exact same speed, then the different movement skills would impose differential temporal demands on you with regard to initiating the batting response – that is, you would likely need to initiate the movement earlier than the professional. At the same time, the difference in movement skills, such as the temporal sequencing of the step and the swing, would require that – even though the exact same kinematic patterns were visually available to both – you and the expert player attune to and use different information (e.g., you may have to rely relatively strongly on early movement kinematics, whereas the expert may even be able to exploit early ball flight information).

Put in more general terms, individual action capabilities fundamentally constrain the way actors encounter their environment, and hence, the requirements differ with respect to the information actors need and attend to. In this respect, attention is an actor's attunement to the information that specifies the relevant properties of the environment–actor interaction. This attunement has multiple characteristics: (a) it is task-specific, (b) it is prone to differ within actors across multiple encounters, and (c) it is also prone to differ among actors (e.g., with different levels of skill). These characteristics follow because individuals guide their (inter)actions by accomplishing certain perceptions (Gibson 1979/1986). Let's elaborate on the characteristics of attunement to further specify our argument: