Burge on the Reach of Perception

1. A central dividing line in the philosophy of perception concerns the demarcation of perceptual capacities from conceptual thought and other higher cognitive processes. On this question, one can broadly distinguish two sorts of position: on the one hand, there are those who favor an *austere* notion of perception, according to which the perceptual faculty is sharply demarcated from higher cognitive functions and outputs representations of only a relatively narrow range of low-level properties such as (in the case of vision) edges, textures, and perhaps colors.[1] On the other hand, there are those who favor a more *liberal* notion of perception, according to which perceptual systems interact in intricate ways with higher cognitive functions and are capable of outputting representations of much more complex properties, perhaps including the properties that distinguish everyday objects such as toasters and teacups.[2]

Tyler Burge's *Origins of Objectivity* constitutes a strong defense of an austere conception of the reach of perception, offering what might well be the most thorough examination of its philosophical and scientific underpinnings to date. Burge's account of perception is fashioned around three familiar general ideas: (1) *Domain specificity*: each perceptual system (for instance, vision) delivers representations of only a relatively small number of low-level attributes (for instance, shape, brightness, color). (2) *Encapsulation*: perceptual systems function relatively independently of input from each other and from higher level cognitive faculties relating to belief and language. Finally, (3) perceptual systems are *shared across species*: for instance, vision functions in the same way, by and large, across most mammalian species (Burge 2010: 101-102).[3]

This demarcation of the perceptual domain entails that representation of more complex objects and properties "depends on capacities that go beyond the perceptual system proper" (ibid.: 101). Accordingly, we do not really perceive everyday objects such as "baseball bats, CD-players,

[1] Cf. Tye 1995; Prinz 2005.
[2] Cf. Bayne 2009; Siegel 2010. A number of important papers on this topic are collected in Hawley and MacPherson, eds., 2011. Note, though, that much of this discussion is carried out in terms of the question of whether these high-level properties are *phenomenally* represented in perception. This question is not central to the line of argument pursued in this paper.
[3] It is worth noting that Burge emphatically does *not* advocate such a view because he believes it constitutes a useful demarcation for theoretical and practical purposes. Instead, Burge's motivation is metaphysical, and he is adamant that his demarcation traces the boundaries of a genuine natural kind, a "cut in nature" (Burge 2010: 9, xii).

hybrid autos," but rather (somehow) reliably come to form beliefs about their presence on the basis of perceiving their low-level properties, such as colors, shapes, and textures.

This conclusion has met with criticism, as one might expect. For instance, Begby (2011) argues that Burge's "peculiarly reductive" account of perception leaves him with little useful to say about how about how perception guides the formation of ordinary empirical beliefs. Critics like Begby may allow that Burge's account gives a correct description of distinctive and important core perceptual capacities, largely innate and shared across species. Nonetheless, they argue that human beings can also come into possession of a wide range of further high(er)-level perceptual attributives. Channeling Thomas Reid's distinction between original and acquired perception,[4] they hold that these high-level perceptual capacities are often *acquired by learning*, and help account for much of the distinctive character of human empirical cognition.

Ned Block has recently come to Burge's defense against such arguments. "Like Burge," he writes, "I am skeptical that we have any 'high level' visual attributives that we do not share at least with our primate cousins" (Block, Ms: 3). Block's strategy is to point out that our phenomenology-based intuitions about particular cases are simply "powerless to distinguish between complexes of colors, shapes and textures and the high level properties that are recognitionally coextensive with them" (Block, Ms: 4). As Block views the dialectic, he is under no pressure to deny the case-based intuitions of those who favor a more liberal account of the reach of perception. Instead, he can simply point out that these intuitions may as well be accounted for in terms of low-level perceptual attributives.

This paper aims to build a new case for a liberal account of the reach of perception. The argument I pursue differs from standard accounts in a number of regards. First, it makes no essential appeal to phenomenology or phenomenal contrast. Second, it builds from the ground up, so to speak, rather than jump directly to intuitions about our perceptual interaction with everyday objects. In both these regards, I believe my arguments will succeed in making better contact with the motivations that underlie the Burgean program. But third, my argument also offers a sketch of an analysis which, if correct, might serve to significantly recast the terms of the debate. This analysis takes its cue from Block's claim, cited above, that intuitions about cases are "powerless

[4] Cf. Reid 1764 VI: 20-23; 1785 II: 21-22.

to distinguish between complexes of colors, shapes and textures and the high level properties that are recognitionally coextensive with them." This claim appears to be guided by what I will call a *presumption of exclusivity*. On Block's view, liberals and austerists are engaged in a debate about the correct level at which to specify the state of a perceptual system at a time. According to the presumption of exclusivity, at most one such specification can be correct, and we have independent reasons (deriving, perhaps, from considerations of scientific parsimony) to favor lower-level specifications where these are available. Accordingly, a liberalist about perception would have to make the case, with respect to particular examples, that not only do we appear to perceive high-level properties (i.e., a familiar face), but that we perceive them *as opposed to* perceiving their corresponding low-level properties (colors, shapes, and textures).

I believe my analysis gives grounds for doubting this presumption: in many cases, it can truly be said that we are perceptually representing both low-level properties *and* high-level properties concurrently. Moreover, in some of the examples I will offer, it will turn out that there is no clear sense in which the high-level properties are "recognitionally coextensive" with any relevant cluster of low-level properties, in the way that Block's argument assumes.

2. In developing my view, I will draw on examples from our perceptual facility in dealing with language. As such, a prefatory note is in order: Burge speaks directly of language perception only once in the entirety of his book. Here is what he has to say:

> In humans and other higher animals, beliefs can affect what is attended to; and attention affects perceptual operations. In humans and higher animals perception interfaces with conception and belief in complex ways. Nevertheless, the processes of perceptual systems, even in humans, are *relatively* independent of higher-order cognitive states. (Language perception is a special case and requires further qualification.) (Burge 2010: 102).

Note that Burge says nothing further to specify in what sense language perception is a "special case" or to cast light on the nature and extent of the "further qualification" required. The proviso appears in the context of discussing the independence of perception from higher cognitive states, so it may be plausible to assume that Burge has in mind well-known effects like the phoneme restoration effect and the McGurk effect. The former offers a striking instance of "cognitive penetration" into the perceptual domain: subjects hear a speech signal as complete even though several phonemes have been removed and replaced with noise; the perceptual system appears to "fill in" the missing phonemes on the basis of projections of semantic cohesion and completeness. By contrast, the McGurk effect illustrates cross-modal interference: subjects are played two video clips in sequence, a first showing a person voicing a particular phoneme (say, /ba/) with matched audio. In the second clip, the audio stays the same (/ba/) but the video is changed to someone uttering a different phoneme (say, /ga/). The visual perception of lips forming the syllable /ga/ appears to override the actual auditory input, leading the subject to perceptually represent a voicing of /ga/.

These examples, interesting and systematic as they appear to be, may nevertheless be compatible with Burge's claim that speech perception is "relatively" encapsulated, that is, largely free from top-down influence and cross-modal interference. I will not pursue the matter here.

But language processing also offers examples that cannot be waved away in this fashion. As has been known for some time,[5] there is a deep mismatch between the acoustic signal and our perceptual representation of speech. Consider, for example, the phenomenon of phoneme grouping, essential to speech processing. Linguists define an "allophone" as the set of acoustically different sounds which speakers of particular languages perceptually group together as instantiating the same phoneme. For example, aspirated and non-aspirated /t/'s are acoustically very different, but are perceptually represented by mono-lingual English speakers as "the same sound," plausibly because the distinction is not semantically relevant in English. In Spanish, by contrast, the distinction between aspirated and non-aspirated /t/'s does correspond to semantically relevant distinctions. Accordingly, they are perceptually represented by first-language Spanish speakers as distinct sounds.[6]

Allophones should qualify as perceptual attributives in Burge's sense. Pertinently, they involve significant constancy mechanisms, grouping distinct physical stimuli as instances of a single kind. But it is evidently a species-specific perceptual capacity. More specifically, there

[5] See, e.g., Fodor, Bever, and Garrett 1974.
[6] See, e.g., Jusczyk 2003; Casserly and Pisoni 2010.

remains a sense in which Burge (and Block) would be right to say that hearing *does* function in the same way across many species, certainly mammalian species. Nonetheless, the study of speech perception is rich with illustrations of the fact that the perceptual capacity of hearing *as deployed toward the specific task of processing speech* has peculiarities unique to our species, peculiarities which ought to be called "perceptual" in their own right.

I hope it will be granted, then, that we have perceptual attributives which we do not share with our "primate cousins." From my point of view, however, a more important question is whether there exist perceptual attributives which are acquired by learning. And it may be argued phoneme grouping does not involve acquired perceptual capacities in this sense: a familiar story from contemporary psycholinguistics has it that we are born with the capacity to discriminate all possible speech sounds deployable in human language, and that we only come to "prune" these abilities with exposure to our first language.[7] If this is correct, then phoneme perception does not support the claim that we can acquire new perceptual forms, however much it may be unique to our species.

With this in mind, our next example will turn to the role of graphemes in reading. Defined on analogy with phonemes in spoken language, graphemes are the minimal units of written language that can make a semantic difference. Each letter of the alphabet is a grapheme, but there are also complex, multi-letter graphemes such as "oi" and "ough" (in English). A study by Rey et al. (2000: 8) offers evidence that "during the acquisition of reading readers develop internal representations of graphemes [resulting from] recurrent associations between orthographic and phonological patterns." Accordingly, "graphemes are processed as perceptual units by the reading system" (ibid.). Similarly, Stanislaus Dehaene, a leading researcher on the topic, writes:

> Our visual system has learned to treat these groups of letters as bona fide units, to the point where we no longer pay attention to their actual letter content. … Our visual system automatically regroups letters into higher-level graphemes, thus making it harder for us to see that groups of letters such as "ea" [e.g., in the context of the word "meat"] actually contain the letter "a"" (Dehaene 2009: 24, 25).

[7] For an overview, see Werker and Tees 2002.

A case can be made, then, that graphemes, not physical marks on the page, are the proper objects of visual perception as deployed toward the task of reading, much like phonemes, rather than acoustic waveforms, are the proper object of auditory perception as deployed toward the task of processing speech. Moreover, the perceptual capacities involved in reading are presumably both species-specific and acquired by learning. (Indeed, as is emphasized in much of the recent literature on grapheme-color synesthesia,[8] the learning process can be immensely difficult, especially in languages with highly irregular phoneme-grapheme correspondence, such as English).

A quick dip into the empirical study of language processing, then, yields the following provisional findings: in addition to showcasing cognitive penetration and cross-modal effects (which may or may not be problematic for Burge, depending on how we cash out the claim that perception is "relatively" encapsulated), language processing displays clear instances of species-specific perceptual capacities that are acquired by learning.

3. Admittedly, phonemes and graphemes are still pretty low-level phenomena, and will hardly whet the appetite of those who favor a more liberal account of the reach of perception. To see how we might go about motivating the transition to richer, higher-level examples, it will be helpful to take a closer look at what happens in reading, and to bring what we find to bear on Block's defensive strategy.

Recall Block's defense of Burge: our "intuitions" about high-level perceptual properties are to be taken seriously, but it will turn out – empirically – that these properties are "recognitionally coextensive" with lower-level perceptual properties, allowing us to say that what we really perceive – strictly speaking – are just "colors, shapes, and textures." Thus, Block saddles the perceptual liberal not just with the burden of making the case that high-level properties are perceptually represented, but also that we perceive them *as opposed to* perceiving the recognitionally coextensive low-level properties.

[8] Cf. Watson, et. al. 2012.

As a general strategy for meeting arguments from phenomenology, this may have its merits. But it is doubtful that it will get any purchase on the example from speech perception: as was pointed out by Jerry Fodor (1983: 54), one has "practically no access to the acoustics of utterances in languages that one speaks." Possibly as a result of a long history of evolutionary adaptation to speech, many low-level features of the speech signal have become completely suppressed or filtered out in perception. Phonemes may be about as low as speech perception goes.[9]

The case of the deployment of the visual system in reading is different, perhaps because reading is a much more recent innovation, scarcely more than 5000 years old: here we do have perceptual access to the physical marks on the page in a way that we do not have perceptual access to the acoustics of speech. Nonetheless, and is suggested by Dehaene's remarks above, what we *perform computations on* when we read are not these marks, but rather a higher-level perceptual object, i.e., graphemes. This insight puts serious strain on what I referred to above as Blocks' presumption of exclusivity: we would be quite wrong to assume that each sensory modality outputs only one perceptual representation per stimulus (even assuming that we can always make sense of the idea of neatly individuated "stimulus packets"). Instead, the deployment of vision in reading strongly suggests that there may be several co-occurring perceptual representations, bringing different properties or property-clusters to computational salience. We are, it seems, perceiving lines on pages and graphemes, side by side. Moreover, it is not generally true that these high-level properties are "recognitionally coextensive" with lower-level properties. In some cases, they might be, but in other cases – as with complex graphemes – they are not. *Reading*, in short, is a very different perceptual exercise than visually tracing marks on the page, although, in a sense, we could not read were we not also capable of concurrently perceiving the lines. One reason why proof-reading, for instance, is such a difficult task is that it requires us to suppress the high-level perception in favor of the low-level, while still keeping faculties related to syntactic parsing and semantic comprehension activated.[10]

---

[9] O'Callaghan 2011 deploys these empirical findings against phenomenal contrast arguments according to which we have perceptual access to the semantic properties of utterances in familiar languages. O'Callaghan's paper is of particular relevance here: while its general strategy is similar to Block's – the phenomenal contrast is real, but can be accounted for in terms of lower-level properties – the lower-level properties he appeals to (i.e., phonemes) are nonetheless lodged at a higher-level than Block general outlook can allow.

[10] As noted above, phenomenology plays no essential role in my argument. With this in mind, it might be worth noting that I am specifically not making a claim about the relative *phenomenal salience* of these concurrent levels of perceptual representation. Instead, my argument is focused on *computational salience* bearing on specific tasks. It

4. Let us turn to seeing how we might deploy these ideas to motivate intuitions about high-level perceptual attributives acquired by learning. My remarks here will be speculative and tentative, but not, I think, entirely without warrant. Here is a familiar case. Let us say I have two ways of telling the time of day. One starts from determining the position of the sun in the sky, the other from looking at a clockface. Starting from the location of the sun, I still need to perform a series of relatively complex calculations to arrive at an estimation of the time. By looking at the clockface, I can tell right away that it is ten to two. It is tempting to say that by looking at the sun, I can infer that it is ten to two; by contrast, looking at the clockface, I perceive that it is ten to two. (Of course, there was a time when extracting the time of day from looking at a clockface did involve inference from more basic perceptual representations. But at some point these inferential steps were dropped and I simply see what time it is. This is part of what we mean by saying that the perceptual attributive in question is acquired by learning.)

Of course, much like in the case of reading, it is plausible that I could not tell the time from the clockface unless I were *also* able to perceive a certain configuration of lines and dots. But much of the information contained in these low-level perceptions is computationally irrelevant to the task at hand. Just as learning how to read involves coming to learn how to construct and manipulate the more abstract perceptual representations, so learning how to tell the time involves learning how to construct and manipulate the more abstract high-level representations that filter out many of the low-level specifics. This is a perceptual capacity, much like the one involved in reading. Contra Block, the fact that we concurrently have perceptual representations of the lower-level properties does not even begin to draw this into doubt. In both cases, we appear to manipulate several layers of representation at once: some specified in terms of low-level properties, others specified in terms of more abstract, schematic properties.

---

may well be that in the battle for phenomenal salience, our perceptual representations at the graphemic level typically loses out to our perceptual representation of lower level features. But it may also be that in another million years or so of adapting evolutionarily to written input, our perceptual system will have gradually suppressed the phenomenological salience of low-level features just as the auditory system has with low-level acoustic features of speech. Be that as it may, I believe appeals to phenomenology does little to settle questions about perceptual representation. In this, I take myself to be in broad agreement with Burge.

Moreover, and again contra Block, it is simply not the case that the high-level properties are always recognitionally coextensive with some particular cluster of low-level properties: there is no unique low-level property specification that is recognitionally coextensive with the high-level perceptual attributive ten-to-two. Learning to perceive the time of day by looking at a clockface involves relevant degrees of desensitization to such information; specifically, it involves the perceptual recognition of a constant element that holds across very different sorts of clockfaces, e.g., Movados and Rolexes. Again, the parallel with reading is instructive: the low-level perceptual capacities involved in reading are highly sensitive to, say, font and color. But these are informationally irrelevant to the computational task at hand, and reading as a perceptual task involves filtering these features out in the construction of a more abstract, higher-level representation.

Once these points sink in then further examples are relatively easy to come by. Common to many of these examples is that they are acquired by learning, often from disciplined and self-conscious inferential strategies exploiting lower-level perceptual capacities. But at some point, the inferential routines are dropped, and we appear to have direct perceptual access to the properties in question. Thus, it takes but little training in music to acquire perceptual recognition of chord gender (major or minor). Similarly, basic training in chess allows you perceptually recognize a particular board configuration as the Sicilian Dragon. Importantly, these representations have the character of what Fodor (1983: 54) dubbed "mandatoriness": once acquired, you cannot perceive harmony as noise, and you cannot but see the board configuration in terms of strategic potentials. Yet for all that, these high-level perceptual representations do not suppress, but co-exist side-by-side with, the lower-level perceptual representations, i.e., the ones that tells us whether the chord is struck on a guitar or a piano, or whether the contrast squares on the board are black or brown.[11]

[11] I note in passing that my argument for multiple concurrent layers of perception also receives broad encouragement from what we know about the functional and anatomical organization of the brain. Crudely put, the brain does not process sensory signals like the digestive system processes food. The digestive system, long-winded and tortuously complex as it may be, is linearly ordered, and there is always some answer to the question of whether last night's dinner is now lodged in, say, the small or the large intestines. In the brain, by contrast, sensory signals can be copied and sent off in multiple directions to be processed in parallel. Notably, there is emerging evidence (cf. McCandliss et al. 2003; Bilalic et al. 2011) that the fusiform gyrus is involved in several visual tasks other than face recognition, such as word form recognition and chess pattern recognition. If this is correct, it is further evidence to suggest that reading and chess playing involves several layers of concurrent perceptual representation.

5. In conclusion, this paper has sought to develop a novel defense of a liberal account of the reach of perception. Drawing on known facts about language processing, I have offered examples of perceptual attributives that are species-specific and acquired by learning. From these examples, I have tried to extract a broader lesson targeted at Block's presumption of exclusivity: contra Block, a single stimulus can give rise to multiple concurrent perceptual representations. Moreover, the properties represented at a higher-level are not always recognitionally coextensive with properties represented at a lower-level. Different concurrent levels of perceptual representation are honed to different computational tasks and isolate distinct properties of the sensory stimulus. Aside from undermining Block's defense of Burge, this argument significantly recasts the terms of the debate: those who favor a liberal account of the reach of perception should not be burdened with making the case that in putative cases of acquired high-level perception, we are perceptually representing high-level properties *as opposed to* low-level properties.

References:

Begby, Endre. "Review of Tyler Burge, *Origins of Objectivity*." *Notre Dame Philosophical Reviews*, February 2011.

Bilalic, Merim, Robert Langner, Rolf Ulrich, and Wolfgang Grodd. 2011. "Many Faces of Expertise: Fusiform Face Area in Chess Experts and Novices." *The Journal of Neuroscience* 31(28): 10206-10214.

Block, Ned. Ms. "What we See." Forthcoming in *Philosophy and Phenomenological Research.*

Burge, Tyler. 2010. *Origins of Objectivity*. Oxford University Press.

Casserly, Elizabeth D. and David B. Pisoni. 2010. "Speech Perception and Production." *Wiley Interdisciplinary Reviews: Cognitive Science* 1(5): 629-647.

Dehaene, Stanislas. 2009. *Reading in the Brain: The Science and Evolution of a Human Invention*. New York: Viking Press.

Fodor, Jerry. 1983. *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: The MIT Press.

Fodor, Jerry, Thomas G. Bever, and Merrill F. Garrett. 1974. *The Psychology of Language: an Introduction to Psycholinguistics and Generative Grammar*. New York: McGraw Hill.

Hawley, Katherine, and Fiona MacPherson, eds. 2011. *The Admissible Contents of Experience*. Oxford: Wiley-Blackwell.

Jusczyk, Peter W. 2003. "The Role of Speech Perception Capacities in Early Language Development." In M.T. Banich and M. Mack, eds., *Mind, Brain, and Language: Multidisciplinary Perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

McCandliss, Bruce D., Laurent Cohen, and Stanislas Dehaene. "The Visual Word Form Area: Expertise for Reading in the Fusiform Gyrus." *Trends in Cognitive Science* 7(7): 293-299.

O'Callaghan, Casey. 2011. "Against Hearing Meanings." *Philosophical Quarterly* 61(245): 783-807.

Prinz, Jesse. 2005. "A Neurofunctional Theory of Consciousness." In Andrew Brook and Kathleen Akins, eds., *Cognition and the Brain: The Philosophy and Neuroscience Movement*, Cambridge University Press.

Reid, Thomas. 1764. *An Inquiry into the Human Mind on the Principles of Common Sense*. In William Hamilton, ed., *The Works of Thomas Reid*, Vol. 1. Edinburgh: MacLachlan and Stewart, 1872.

Reid, Thomas. 1785. *Essays on the Intellectual Powers of Man*. In William Hamilton, ed., *The Works of Thomas Reid*, Vol. 1. Edinburgh: MacLachlan and Stewart, 1872.

Rey, Arnaud, Johannes C. Ziegler, and Arthur M. Jacobs. 2000. "Graphemes are Perceptual Reading Units." *Cognition* 75(1): B1-B12.

Siegel, Susanna. 2010. *The Contents of Visual Experience.* Oxford University Press.

Tye, Michael. 1995. *Ten Problems of Consciousness*. Cambridge, MA: MIT Press.

Van Gulick, Robert. 1994. "Deficit Studies and the Function of Phenomenal Consciousness." In G. Graham and G. Lynn Stephens, eds., *Philosophical Psychopathology*, Cambridge, MA: The MIT Press.

Watson, Marcus R., Mark R. Blair, Pavel Kozik, Kathleen A. Akins, and James T. Enns. 2012. "Grapheme-Color Synesthesia Benefits Rule-Based Category Learning." *Consciousness and Cognition* 21(3): 1533-1540.

Werker, Janet F., and Richard C. Tees. 2002. "Cross-Language Speech Perception: Evidence for Perceptual Reorganization During the First Year of Life." *Infant Behavior and Development* 25(1): 121-135.